IAAP

BASES PARA LA
EVALUACIÓN DE IMPACTO
DE POLÍTICAS PÚBLICAS

Evaluación Políticas Públicas

Instituto Andaluz de Administración Pública
CONSEJERÍA DE LA PRESIDENCIA, ADMINISTRACIÓN PÚBLICA E INTERIOR

CRÉDITOS

AUTOR PRINCIPAL

Eduardo Serrano Padial. Consejería de Agricultura, Ganadería, Pesca y Desarrollo Sostenible

COLABORADORAS

Adolfina Martínez Guirado. Consejería de Empleo, Empresa y Comercio
Juana M.ª García Simón. Consejería de Economía y Conocimiento
Ana Díaz Gil. Oficina Técnica de EPP. Instituto Andaluz de Administración Pública

COORDINADORA

María del Mar Herrera Menchén. Instituto Andaluz de Administración Pública

EDITA

Instituto Andaluz de Administración Pública. Junta de Andalucía. 2019

IMPRIME

Servicio de Publicaciones y BOJA

CÍTESE

Serrano, E., Herrera, M., Martínez, A. y Simón, J. y Díaz, A. Bases para la Evaluación de Impacto de Políticas Públicas. IAAP: Sevilla

Depósito legal: SE-543-2019

ecoedición

Tinta sin metales pesados y papeles procedentes de una gestión forestal sostenible

Impacto ambiental	Agotamiento de recursos fósiles	Huella de carbono
por producto impreso	0,35 kg petróleo eq	1,14 Kg CO_2 eq
por 100 g de producto	0,12 kg petróleo eq	0,41 Kg CO_2 eq
% medio de un ciudadano europeo por día	7,85 %	3,72 %

JUNTA DE ANDALUCIA
CONSEJERÍA DE MEDIO AMBIENTE Y ORDENACIÓN DEL TERRITORIO

reg. n.º: 2019/44

Más información en www.ecoedicion.eu

ÍNDICE

ACLARACIONES METODOLÓGICAS 8

INTRODUCCIÓN 9

CAPÍTULO 1: UNA TAREA COMPLICADA: DEFINIR EL IMPACTO 11

- ASPECTOS BÁSICOS DE LA EVALUACIÓN DE IMPACTO 11
- TIPOS DE EVALUACIÓN DE IMPACTO SEGÚN EL MOMENTO EN EL QUE SE REALIZAN 13
- ELEMENTOS CLAVE: LA CAUSALIDAD Y EL CONTRAFACTUAL 16
  - Causalidad 16
  - Contrafactual 17
- UN DEBATE IMPORTANTE: ATRIBUCIÓN Y CONTRIBUCIÓN 19
- DEJANDO LAS COSAS CLARAS: RESULTADOS vs IMPACTO 20
- DEFINICIÓN DE EVALUACIÓN DE IMPACTO 23
- ¿CUÁNDO MEDIR EL IMPACTO? 24
  - Criterios relacionados con el diseño y la implementación del programa 24
  - Criterios relacionados con el análisis de los efectos y el contrafactual 25
  - Criterios relacionados con la propia política pública 25

CAPÍTULO 2: EN BUSCA DE LA CAUSALIDAD: EL CONTEXTO COMO PUNTO DE PARTIDA Y LA LÓGICA DEL PROGRAMA 27

- ANÁLISIS CAUSAL: LA LÓGICA DE LA ACTUACIÓN 28
  - El Enfoque del Marco Lógico 29
  - La Teoría del Programa 30
  - El Análisis Causal 31
- LA LÍNEA BASE: DESCRIBIENDO EL CONTEXTO 33
  - Diseccionando los Objetivos Estratégicos 33
  - Selección de indicadores de la Línea Base 35

CAPÍTULO 3: MIDIENDO EL IMPACTO: DEFINIR EL CONTRAFACTUAL 39

- EL CONTRAFACTUAL O COMO VALORAR LOS EFECTOS DEL NO PROGRAMA 40
  - Validez de los resultados 40
- CONTRAFACTUALES CON GRUPOS DE CONTROL 41
  - Experimentales o aleatorios 41
  - Grupos reconstruidos: Emparejamiento o Matching 42
  - Dobles Diferencias 44

CONTRAFACTUALES REFLEXIVOS – SIN GRUPOS DE CONTROL 46
Series temporales 46
Estudios de caso 48

CAPITULO 4: UNA PROPUESTA METODOLÓGICA A MODO DE CONCLUSIÓN ... 51

PARA SABER MÁS: 55

Características de los indicadores de contexto 58
Proyectando indicadores: La Línea Base de la Estrategia de Impulso del Sector TIC Andalucía 2020 61
La Teoría del Cambio: Aplicación a un programa educativo 63
Ejemplos de contrafactual en políticas sociales y económicas 68

BIBLIOGRAFIA 79

BASES PARA LA

EVALUACIÓN DE IMPACTO

DE POLÍTICAS PÚBLICAS

ACLARACIONES METODOLÓGICAS

Como ya se señala en la introducción, este documento pretende ser sobre todo divulgativo. Con él se busca acercar los conceptos y las metodologías de las evaluaciones de impacto a aquellas personas que, sin tener una formación o un conocimiento profundos, están implicadas de algún modo en la evaluación de programas públicos. Precisamente por este carácter divulgativo y de primer acercamiento a la materia, se han simplificado, consciente e intencionadamente, los conceptos y metodologías que se recogen en el texto. Sin embargo, entendemos que es necesario aclarar este punto para evitar confusiones posteriores cuando deban abordarse las evaluaciones.

El uso que se hace en el texto de "Población" y de "Contexto"es equivalente. Aunque se entiende que el contexto es un concepto más amplio que población, en el sentido de que la segunda hace referencia exclusivamente a un conjunto de individuos, en aras de la simplicidad se ha optado a usarlas como sinónimos en algunas partes del texto.

Otro aspecto cuyo uso puede resultar extraño a quien haya trabajado con él, es el de la "Línea Base". En la metodología que se propone, la "Línea Base" es un conjunto limitado de indicadores fundamentales que permite valorar los cambios en la población o el contexto de la actuación pública. Por lo tanto, no debe confundirse con otros significados que a veces se usan bajo este nombre o nombres similares; compendio de indicadores que definen pormenorizadamente el contexto de la actuación, conjunto de condiciones previas que deben cumplirse para llevar a cabo la actuación, etc. Por otro lado, nos referiremos a estos indicadores como **Indicadores de contexto**[1], aunque, en la terminología habitual se pueden entender también como **Indicadores de impacto**[2].

Finalmente, se ha optado por el uso generalizado de los términos **Programa y Acción pública** para señalar cualquier tipo de programación formal (plan, proyecto, etc.) y para indicar, de manera general, una actuación o política pública, respectivamente.

---

1. Los indicadores de contexto "proporcionan información sobre elementos socioeconómicos u otros aspectos relevantes de la situación de partida del Plan" (Manual de elaboración de Planes Estratégicos de Políticas Públicas en la Junta de Andalucía. IAAP 2017.

2. Los indicadores de impacto no son más que indicadores de contexto que sirven para medir el impacto.

INTRODUCCIÓN

Quizá no haya cuestión más transcendente para una actuación pública que preguntarse si **ha servido para algo** y, en caso afirmativo, para qué, en qué ha influido, qué se ha conseguido tras su puesta en marcha.

A esta pregunta central pretende dar respuesta la **Evaluación de Impacto**, un tipo de evaluación cuyo único objetivo es conocer en qué medida un plan, programa o proyecto ha influido en la población o en el contexto de las actuaciones, causando los efectos que se perseguían con él.

En el año 2016, el Gobierno Andaluz inició un proyecto para desarrollar la evaluación de políticas públicas en la Junta de Andalucía. Entre otras actividades, esta iniciativa puso en marcha un trabajo colaborativo mediante la creación de un grupo de personas con experiencia y/o conocimientos en la materia que, de manera voluntaria, propusiera y abordara conjuntamente las diferentes tareas que esta empresa requiriera. Entre las tareas que se han ido abordando, se encuentra la elaboración de una metodología propia, sencilla y práctica, para la planificación y evaluación de las políticas públicas.

De esta forma, y después de abordar otros aspectos fundamentales como la ***planificación***, la ***evaluación ex-ante*** o la misma ***evaluabilidad de los programas públicos***, ha llegado el momento de abordar otras evaluaciones que completaran el ciclo de las políticas públicas, más allá de las fases de diseño e implementación, incluyendo la evaluación de los efectos de las actuaciones puestas en marcha.[3]

La evaluación de impacto que aquí se plantea, se enmarca en ese concepto de los ***efectos causados*** por el plan, analizados ***tras su implementación*** y en el momento adecuado para que se hayan podido producir. En definitiva, conocer ***los cambios en la población o en el contexto*** que han sido fruto, de algún modo, de la política pública.

Al mismo tiempo, se opta por unas determinadas metodologías para abordar este tipo de evaluación de políticas públicas, de cara a contar con una forma común de emprender este proceso para toda la Junta de Andalucía. La elección de métodos y la forma en que se proponen, responden al ***enfoque eminentemente práctico*** que se le quiere dar a estas guías, por lo que, lejos de aproximaciones teóricas, las propuestas recogidas en los capítulos siguientes y en el apartado “Para saber más”, que pueden resultar eclécticas desde el punto de vista de la teoría y la conceptualización de la evaluación, han sido propuestas a través del debate y el consenso, de manera que respondan a las necesidades y la visión propias del equipo de evaluación que luego tendrá que usarlas.

3. A los efectos de este manual se usarán como sinónimos “planificación pública” y “política pública”. De igual modo, se entenderá que el término “planificación pública” comprende los conceptos de “política pública”, “plan estratégico”, y “programa público”. Por su parte, se usarán como sinónimos los términos de “acción pública” e “intervención pública”.

01

CAPÍTULO 1: UNA TAREA COMPLICADA: DEFINIR EL IMPACTO

ASPECTOS BÁSICOS DE LA EVALUACIÓN DE IMPACTO

El concepto de impacto no siempre es tan evidente como pueda parecer. Para ilustrar esta afirmación imaginemos un programa que, por ejemplo, va dirigido a resolver un problema de bajo poder adquisitivo detectado en un sector de la población.

En principio, podríamos pensar que si, tras la puesta en marcha del programa, la población objetivo ha incrementado su renta, el programa habría tenido un impacto positivo. Pero ¿realmente sería así? Vamos a profundizar un poco más en ello.

Primero, deberíamos tener claro si el incremento de renta ha supuesto un incremento del poder adquisitivo. Aunque esta cuestión no es específica de la evaluación de impacto, nos ayuda a entender qué **debemos atender el objetivo que se persigue** y evitar dar por buenos unos efectos que, si bien parecen apuntar en la buena dirección, no terminan de mostrarnos si se ha logrado la meta perseguida. En este caso concreto, podemos encontrarnos que el coste de la vida también ha subido y que realmente no se ha producido la mejora que esperábamos.

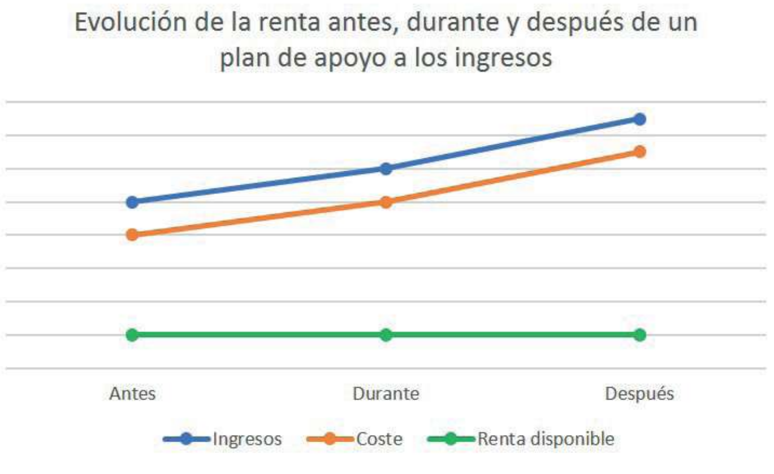


Figura 1: Enfoque incorrecto del efecto de un plan, los ingresos se han incrementado pero la renta disponible no ha sufrido variaciones

En segundo lugar, debemos tener en cuenta **la influencia del tiempo en los efectos**. No todos los efectos se muestran de manera inmediata o, incluso, pueden evolucionar con el tiempo, por lo que debemos tener la seguridad de que la evaluación se afronta en el momento adecuado que nos permita poder valorar los efectos, cuando realmente se han podido producir. Siguiendo con el ejemplo anterior, podríamos encontrarnos que hasta pasado un tiempo desde la ejecución del programa, no se muestran los efectos sobre la renta.

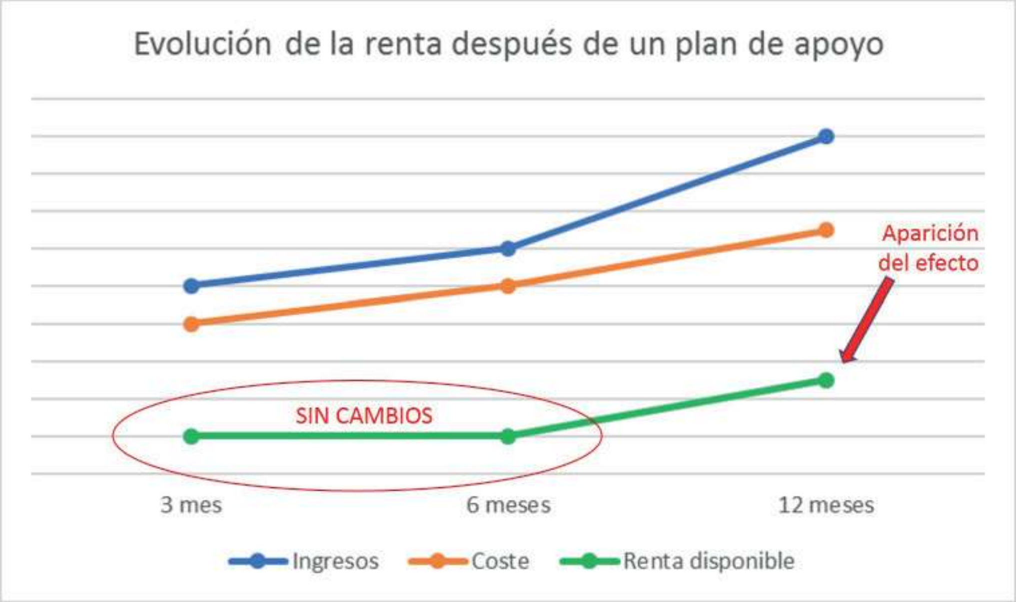


Figura 2: Influencia del tiempo en la aparición de los efectos de una intervención

En general, **es esperable que los efectos surjan después de la actuación**, de modo que, aunque en algunos casos puede ser necesario o recomendable medir resultados durante la ejecución del propio plan o programa, por defecto se suelen plantear **evaluaciones ex-post**. Idealmente, y en la medida en que sea posible, debería contemplarse la **temporalidad de los efectos** en el propio plan, de tal manera que cuando se diseñe la actuación, ésta recoja orientaciones acerca de cuándo afrontar la evaluación de impacto.

"En este sentido, se trata de determinar cuál es el momento más relevante para hacerse la pregunta: «¿Qué ha pasado en comparación con lo que habría ocurrido si la intervención no se hubiera puesto en práctica?»"

(BLASCO, J. y CASADO, D. 2009: 18)

Por último, deberemos enfrentar la cuestión de **qué ha pasado con el resto de la población que no ha sido atendida por el programa**. ¿Hay alguna diferencia entre ella y la población objetivo? ¿La renta o el poder adquisitivo de estas poblaciones se han incrementado por igual? ¿Hay diferencias significativas? En definitiva, ¿Podemos afirmar que el programa ha colaborado al incremento de la renta o, por contra, esto hubiera ocurrido con independencia de él?

Podemos sintetizar los elementos a tener en cuenta para abordar el estudio de impacto de una actuación, en el siguiente cuadro:

ASPECTOS BÁSICOS DE LA EVALUACIÓN DE IMPACTO DE UN PLAN ESTRATÉGICO

Pertinencia: El impacto o efecto esperado relacionado con los objetivos del plan que se va a evaluar. Aunque cualquier actuación puede tener efectos no programados previamente, se buscará conocer hasta qué punto se ha logrado influir en el contexto de la misma, conforme a los objetivos estratégicos establecidos.

Temporalidad: La elección del momento de la evaluación debe ser tal que los efectos puedan haberse producido. En la medida de lo posible, este momento debería recogerse en el sistema de evaluación del plan o programa.

Causalidad: Los efectos deben tener su origen en las actuaciones del plan. Este aspecto es clave a la hora de entender y evaluar el impacto. ¿Los efectos son consecuencia de las actuaciones o medidas puestas en marcha por el plan que se evalúa? ¿Existe una relación causa-efecto entre los resultados del plan y los impactos observados?

Todas estas cuestiones no son fáciles de afrontar y mucho menos de responder y, de algún modo, explicitan la dificultad que entraña la evaluación de impacto.

TIPOS DE EVALUACIÓN DE IMPACTO SEGÚN EL MOMENTO EN EL QUE SE REALIZAN

Otro aspecto esencial de la evaluación de impacto es el momento en el que se evaluarán los efectos, ya que ello marca una diferencia importante en cuanto al tipo de evaluación de impacto que se va a abordar. Básicamente existen **dos tipos de evaluación de impacto** en función del momento en que se realizan: evaluación de impacto **ex-ante** y **ex-post**. Tomando como referencia la definición de la OCDE (2014):

> ***Evaluación de impacto ex-ante**: Implica realizar un análisis prospectivo de cuál podría ser el impacto de una intervención, a fin de apoyar a la formulación de la acción pública. Este tipo de evaluación formaría parte de la actividad de análisis y planificación del ciclo de la política y, por lo tanto, es previo (ex-ante) a su implementación.*

Este concepto de evaluación de impacto ha sido recientemente afinado por la Unión Europea. La UE recoge en sus reglas de buen gobierno o Better Regulation[4] la realización de este tipo de evaluaciones, pero ya no las nombra evaluaciones de impacto, sino que emplea el término **Inception Assessment**, algo así como "Evaluación Previa" o "Evaluación Inicial". Esta acepción se ajusta más al objetivo de este tipo de análisis, que es prever los efectos de una determinada actuación, **valorando los posibles escenarios** que pueden darse con sus resultados e impactos previsibles. De esta manera, este tipo de evaluación inicial obedece más al análisis del potencial impacto de diferentes alternativas de acción, a fin de poder seleccionar entre ellas.

Figura 3: Evaluación de impacto inicial o *Inception assessment*

Sin embargo, existe otro enfoque para la evaluación del impacto ex-ante, que sí pretende establecer el impacto de una determinada política para **prever medidas correctoras o para prevenirlo** rediseñando la actuación. Es el caso de las **evaluaciones sobre elementos transversales**, como pueden ser las de impacto ambiental, las de impacto de género o las de impacto sobre la salud. En estos casos, no se realiza la evaluación para decidir si se implementa o no una determinada política y en qué sentido, sino que una vez tomada la decisión y definida su forma de implementación, se valoran los efectos sobre el medio ambiente, la igualdad de género o la salud pública, comprobando si se cumplen los objetivos transversales en estos aspectos y, a partir de estos datos, incluir medidas correctoras dentro de la propia actuación o, incluso, rediseñar aquellos aspectos de la actuación que tienen efectos no deseados.

4. Las guías de la Comisión Europea sobre mejor gobierno se encuentran en el sitio https://ec.europa.eu/info/law/law-making-process/planning-and-proposing-law/better-regulation-why-and-how/better-regulation-guidelines-and-toolbox_en (visto en julio de 2018)

Figura 4: Evaluación de impacto transversal

De esta forma, tendremos que la Evaluación Previa/Inicial (Inception Assessment) ayudará a decidir la manera de abordar la consecución de los objetivos de la estrategia, y la Evaluación de Impacto Transversal ayudará a adecuarla a los objetivos transversales o a las normas que condicionan la actuación, evitando o minimizando efectos negativos y fomentando efectos positivos relacionados con ellos.

Dada la naturaleza de ambos tipos de evaluación de impacto, éstas pueden ser simultáneas para un mismo programa. De manera gráfica, las dos evaluaciones de impacto ex-ante con las que podemos encontrarnos con mayor frecuencia generarían dos vías de modificación de los elementos del programa evaluado:

Figura 5: Tipos de evaluaciones de impacto *ex-ante* de un programa

Veamos ahora el otro tipo de evaluación de impacto:

> **Evaluación de impacto ex-post***: Este tipo de evaluación pretende comprender en qué medida y cómo, una intervención ha contribuido a resolver el problema que pretendía abordar. La evaluación de impacto se centra en los efectos de la acción pública y forma parte de la gestión del ciclo de políticas públicas, permitiendo tener en cuenta estos efectos y entender los mecanismos por los cuales se han producido de cara al diseño de futuras políticas.*

Mientras que otro tipo de evaluaciones ex-post pueden cubrir una gama más amplia de criterios de evaluación, tales como la adecuación del diseño de la actuación, la eficacia y la eficiencia, etc.

La evaluación de impacto nos ofrece información específica sobre los **efectos de la acción pública sobre el contexto al que se dirige**, y sobre la **contribución** de estos efectos a la **consecución de los objetivos** de la misma.

ELEMENTOS CLAVE: LA CAUSALIDAD Y EL CONTRAFACTUAL

> *"Las evaluaciones de impacto constituyen un tipo particular de evaluación que pretende responder a una pregunta específica de causa y efecto:* ***¿Cuál es el impacto (o efecto causal) de un programa en un resultado de interés?*** *Esta pregunta básica incorpora una dimensión causal importante."*
>
> (Gertlery otros. 2017:9)

Este es, por lo tanto, el rasgo definitorio de este tipo de evaluaciones, **la relación entre la actuación y su efecto en el contexto en el que se aplica**, por lo que, para poder medir el impacto, lo primero que debemos analizar es la **causalidad**.

Por otro lado, mediante este tipo de evaluaciones queremos saber el efecto que se le puede atribuir a la actuación evaluada; es decir, conocer la diferencia de efectos entre el programa y el "no-programa" entre la población beneficiada y la que no ha sido partícipe de la actuación. Esta última situación es lo que se entiende por **contrafactual**.

Causalidad

Como señalábamos, la causalidad explica la **relación entre las actuaciones y sus efectos, o dicho más sencillo, los efectos sobre el contexto o la población** son debidos a las actuaciones de un plan estratégico que incide sobre ellos. Esto evita que se asignen efectos que realmente no se deben a una acción determinada o distinguir relaciones espurias producto de casualidades o falsas correlaciones[5].

5. Por ejemplo, un incremento de la producción agraria que coincide con un nuevo programa de formación para agricultores puede llevar a relacionar ambas cuestiones y que, sin embargo, éste incremento sea realmente debido a una mejora en la meteorología.

Aunque hay casos en los que no se puede establecer una relación causal pormenorizada, el análisis causal está implícito en las principales propuestas metodológicas de diseño y planificación estratégica, desde el Enfoque del Marco Lógico (EML) hasta la Teoría del cambio[6].

Cualquier programa elaborado con estas metodologías, nos mostrará la lógica del mismo y sus relaciones causales que servirán de base para la posterior evaluación de impacto. Dicha evaluación, además, permitirá corroborar si dichas relaciones, planteadas a priori en el plan, han sido las que efectivamente han operado para conseguir los efectos observados.

En la siguiente figura se muestra como ejemplo las relaciones causales dentro del Enfoque del Marco Lógico:

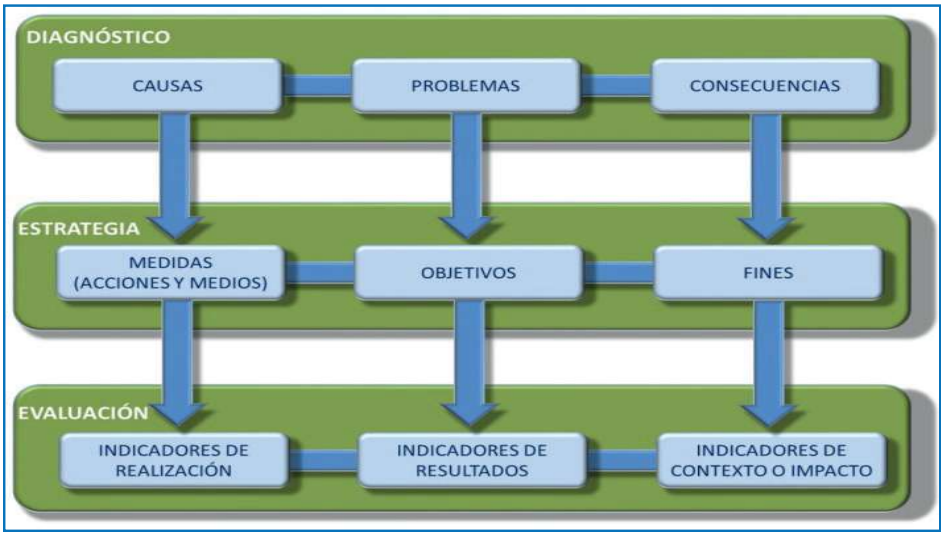


Figura 6: Relaciones entre elementos del Marco Lógico para la planificación y la evaluación

Contrafactual

Para la determinación de la posible relación causal entre "programa" y "efecto", no basta la existencia de una **relación lógica** entre ambos, sino que también deberemos valorar si realmente **los cambios producidos han sido fruto de la aplicación** del programa o, por el contrario, si dichos cambios son el **resultado de la evolución propia del contexto**, de manera que hubieran ocurrido tanto si se ejecuta el programa como si no hubiera existido.

Consecuentemente, para establecer la causalidad debemos analizar dos cuestiones: la **lógica de la actuación y el cambio operado en el contexto, o en la población, respecto a la hipotética situación de que no se hubiera producido dicha actuación**. Esa situación hipotética se conoce como el contrafactual del programa, es decir, el "grupo

6. Ver el Manual de planificación estratégica, así como el de Evaluación ex-ante del IAAP (IAAP: 2017).

de control" o referente que nos sirve para poder establecer si el programa evaluado ha supuesto realmente una diferencia, ya que es el grupo que no ha sido beneficiario del programa pero que contaba con las mismas características de partida. De alguna manera, estamos tratando de determinar qué hubiera pasado a ese grupo o contexto si no se hubiera realizado el programa.

Pero, veámoslo con un ejemplo sencillo. Supongamos que se ha ejecutado un plan especial de incorporación de jóvenes a su primer empleo tras el cual 200 jóvenes han encontrado su primer trabajo. Este resultado por sí mismo no nos dice si el programa ha sido el causante de estas incorporaciones o si es fruto de una situación económica favorable. Para ello, deberíamos saber cuántos de los jóvenes habrían encontrado su primer trabajo en el mismo periodo si no se hubiera implementado ese plan especial (en nuestro ejemplo, 50). Este escenario, el "No-Programa", será nuestro contrafactual y, en el ejemplo, nos indica que el impacto real del plan no serían 200 jóvenes trabajando por primera vez sino 150, ya que 50 habrían encontrado su primer trabajo aún sin el programa de inserción.

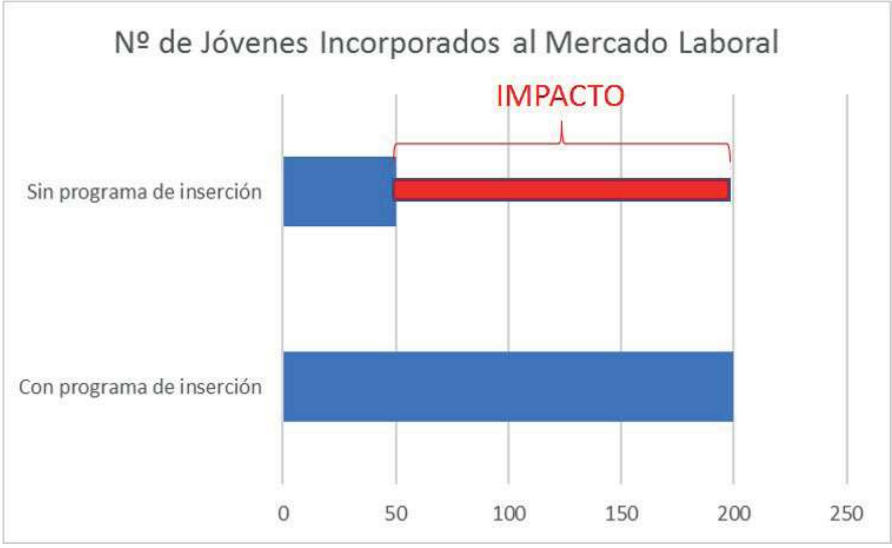


Figura 7: Impacto en la incorporación al mercado laboral de un programa de inserción laboral de jóvenes

Por lo tanto, para poder evaluar realmente si un programa ha tenido impacto, tan importante es la lógica interna del programa, la relación causa-efecto entre los elementos del mismo (Problemas-Objetivos-Medidas) como el análisis del contrafactual, ya que sin él no podremos llegar a valorar los efectos del plan tras su implementación.

Es decir, la evaluación de impacto permite determinar si los efectos han sido producidos por el programa o, en alguna medida, el programa ha contribuido a que los efectos sean mayores a los que se hubieran experimentado si no hubiera existido el mismo.

UN DEBATE IMPORTANTE: ATRIBUCIÓN Y CONTRIBUCIÓN

Relacionado con la causalidad, existe una cuestión que condiciona qué considerar como impacto o efecto, debido al plan o programa que se ha implementado. Esta cuestión es **¿En qué medida se pueden atribuir los cambios observados a la actuación llevada a cabo?** Este es el debate entre "Atribución" y "Contribución".

UN DEBATE DELICADO

Atribución: El impacto o efecto se debe **exclusivamente** a los resultados del plan. Para determinar la atribución de una actuación debemos poder aislar el contexto de la misma y compararlo con contextos iguales no afectados por dicha actuación.

Contribución: Los resultados de la actuación **han colaborado** a conseguir los efectos detectados, pero no son su única causa. En este caso no podemos afirmar que dichos efectos se deban exclusivamente al plan o programa llevado a cabo, aunque se debe poder relacionar con ellos.

El Banco Mundial en su manual (Gertler y otros, 2017:4), señala que la **atribución** es la característica definitoria de la evaluación de impacto, y establecer la relación causal entre el proyecto y los efectos obtenidos es clave, ya que generalmente son difíciles de llevar a cabo, necesitando datos muy precisos o técnicas experimentales. De esta manera, salvo que contemos con datos muy fiables de la población no beneficiada por el programa, que permitan establecer una relación robusta y veraz entre los efectos y el programa evaluado, siempre podrá quedar en entredicho la atribución de dichos efectos a éste e, incluso en este caso, siempre quedará una sombra de duda.

Por otro lado, no debemos olvidar que para muchas de las actuaciones a las que nos vamos a enfrentar, a la hora de evaluar el impacto, no podremos llegar a tal grado de precisión. En algunos casos, será por el simple hecho de que son actuaciones pasadas que no permiten un análisis del contrafactual de este tipo, en otros por la propia naturaleza de la población o de los objetivos, que no permiten tal grado de precisión. Por ejemplo, un programa que se plantea una nutrición más sana y equilibrada tendrá efectos positivos en la salud de la población beneficiaria. Sin embargo, la cantidad de factores que pueden influir en los indicadores (higiene, actividad física, riesgos laborales, estilos de vida previos, etc.) es tal, que sería imposible afirmar que la mejora sea atribuible exclusivamente a la mejor nutrición que ha promovido el programa. No obstante, sí podríamos evaluar dicho programa para afirmar **si ha contribuido** a una mejor salud de la población beneficiaria.

Otra cuestión es el **ámbito limitado de la actuación**, de forma que, si sobre un problema actúan varios factores y solamente se abordan algunos de

ellos, ha de descontarse el efecto de los factores no atendidos. Volviendo al ejemplo de la inserción de jóvenes, el plan especial tendría como objetivo incrementar la inserción laboral de los jóvenes, pero puede haber abordado solamente algunos elementos que la hayan facilitado (p. Ej. la formación para el mercado laboral) y, sin embargo, sería difícil atribuir la inserción conseguida exclusivamente a las mejoras introducidas por el programa (en este caso, mejora de la formación).

En línea con lo anterior, nos encontramos en un mundo cada vez más complejo, con mayores interrelaciones a nivel global que pueden afectar los resultados y los efectos de un programa, sin que sea posible prever durante el proceso de formulación del plan todas las hipótesis o condicionantes que puedan producirse. Por ejemplo, hasta qué punto podríamos aislar el efecto de un programa de promoción turística vía web, sobre el incremento de pernoctaciones en los hoteles andaluces del resto de factores (recomendaciones personales, actividad de los departamentos comerciales, economía colaborativa, conflictividad en zonas competidoras o influencia de series de televisión) como para poder atribuirle dicho incremento[7].

Aunque es tentador querer tener la certeza y poder afirmar que el plan o política pública es la causante en exclusiva de los efectos causados, sobre todo si se consideran exitosos, debemos tener presente las limitaciones de las técnicas de evaluación y su coste-beneficio, dado el esfuerzo de tiempo y recursos que requiere.

En este contexto, pretender atribuir un determinado efecto de manera exclusiva puede ser una carga excesiva, cuyas conclusiones no merezcan el esfuerzo que requiere. Por esto, y con el objeto de poder realizar evaluaciones de impacto que, al menos, nos orienten sobre los posibles efectos de un programa, dejaremos de lado este debate, de forma que ***las metodologías que se proponen en esta guía implican más un análisis de la contribución del plan, entendida como la relación causal, aunque no exclusiva, entre el programa y una aproximación de los cambios en la población beneficiaria o en el contexto, respecto a la hipotética situación de que no se hubiera hecho nada***. Si de dicho análisis se concluyera la necesidad de profundizar en los detalles de dicha relación y de los efectos detectados, sería el momento de valorar la posibilidad de analizar la atribución del programa en cuestión.

DEJANDO LAS COSAS CLARAS: RESULTADOS vs IMPACTO

El último aspecto que nos queda por aclarar es la diferencia entre "Impacto" y "Resultado" de un programa. Aunque a priori pudiera parecer que son conceptos fácilmente diferenciables, nos podemos encontrar que, conforme abordamos las evaluaciones, la frontera entre ambos se diluye haciendo más probable la confusión entre ambos.

La primera diferencia es una cuestión de nivel o jerarquía en la planificación. **Los "Resultados" son la consecuencia del conjunto de productos gene-**

7. Como ejemplo de esta reflexión en el ámbito del márquetin online se recomienda esta breve lectura https://www.divisadero.es/de-la-atribucion-a-la-contribucion/ (visitado en julio de 2018)

rados por las medidas concretas del programa, y colaboran a cumplir con los **objetivos más concretos de la programación** (objetivos específicos o generales según el caso).

Por otra parte, **el Impacto son los efectos que ha causado el programa sobre el contexto al que se dirige como fruto de sus Resultados, y que se alinean con los grandes objetivos de la política pública** (objetivos estratégicos).

En la cadena de resultados "tipo" de cualquier política pública, como la que se muestra en la figura, se puede contemplar la secuencia lógica de productos, resultados y efectos y los procesos de evaluación asociados a cada uno.

Figura 8: Cadena de resultados y tipos de evaluación de una política pública (Feria Moreno, A: 2017)

La segunda diferencia, y quizá la más importante, es la conceptual. Los **"Resultados" son fruto inequívoco de la implementación de la política pública**, hasta cierto punto puede afirmarse que son el producto final de la misma. Sin embargo,**"Impacto" son los efectos que el programa** ha causado en la **población**. Si la población que no ha participado en el programa (el grupo contrafactual) ha sufrido los mismos efectos, no podremos asignarle a dicho programa la causa de los mismos.

En definitiva, un programa generalmente tendrá resultados y productos, pero estos no necesariamente tendrán impacto en las situaciones de partida.

Para ilustrar estas diferencias retomemos el ejemplo sobre el programa de mejora de rentas del principio de este capítulo. El programa tendría como objetivo la mejora de la renta disponible y para ello se ha optado por unas actuaciones que permitan incrementar los ingresos.

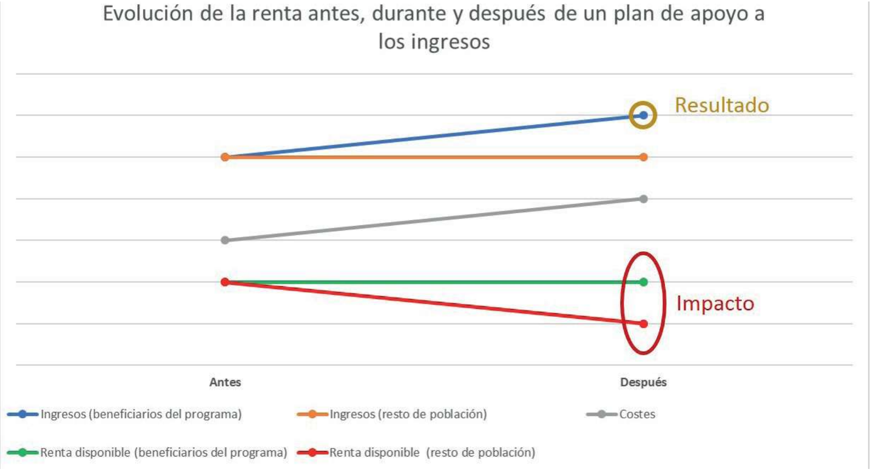


Figura 9: Resultado e impacto de un plan de apoyo a los ingresos

Como puede verse en la gráfica, el programa ha conseguido elevar los ingresos de los participantes (Resultado) pero no ha conseguido elevar la renta debido a un incremento paralelo de los costes, por lo que, en este caso, el efecto sobre la renta disponible sería nulo y no habría Impacto. Sin embargo, si prestamos atención al comportamiento de los ingresos de la población no beneficiaria, vemos que ha sufrido una congelación de ingresos que le ha llevado a una pérdida efectiva de la renta. En este sentido, el programa sí ha causado un impacto, no el esperado de incrementar la renta, pero si ha supuesto un diferencial de renta para quienes han participado en él, permitiendo mantenerla.

Si nos centramos en la parte de los ingresos, se puede comprobar la dificultad que a veces supone distinguir entre "Resultados" e "Impactos". Se observa que hay un diferencial entre los ingresos de la población beneficiaria y el resto, y podríamos pensar que aquí habría un impacto. Sin embargo, es realmente un resultado; el objetivo específico es incrementar los ingresos (lógicamente de la población beneficiaria) y eso se ha logrado, pero el verdadero impacto ha sido el conseguir mantener la renta disponible a través del programa, frente a la pérdida de renta de la no beneficiada por la política pública.

RESULTADOS VS IMPACTO ¿QUÉ MEDIR?

Todo plan de evaluación que se preste incluirá siempre la medida de los resultados de un programa, al igual que sus productos. Sin embargo, no siempre se podrá o merecerá la pena medir los impactos, como se señala en el último apartado de este capítulo. Para poder tomar estas decisiones es imprescindible diferenciar claramente entre ambos.

El impacto implica que sea el efecto esperado y previsto por el objetivo, estando contemplado en la lógica de actuación. Si, como en el ejemplo, se detectan efectos no esperados, estos requerirán un tratamiento aparte

DEFINICIÓN DE EVALUACIÓN DE IMPACTO

Aunque hay numerosas definiciones de impacto que podemos encontrar en la bibliografía especializada, de cara a esta guía, y siguiendo lo que acabamos de exponer, podemos definir el **impacto** como:

> *El efecto de una determinada actuación sobre la población que ha participado en ella o sobre el contexto sobre el que se ha aplicado conforme a lo establecido en los objetivos de dicha actuación.*

Esta definición plantea tres cuestiones fundamentales:

Primero, debe haber un **contexto o situación de partida**. Este contexto vendrá **definido por una serie de indicadores de referencia** que constituirán lo que más adelante definiremos como **Línea Base**.

Segundo, debe **haberse producido un Efecto** (incluso un no efecto) lo que implica una dimensión temporal; debe haber transcurrido el **tiempo necesario** para que se pueda producir.

Y, tercero, el efecto debe ser fruto del programa, lo que supone una causalidad entendida en dos sentidos; por un lado, que **exista una lógica de la actuación** que establezca la relación entre la actuación y el efecto conseguido y, por otro, que **se pueda establecer al menos la contribución** del programa a dicho efecto, comparando la situación de la población beneficiaria con una población o modelo de referencia, que represente la situación derivada de la no participación en el mismo.

A lo largo de los dos siguientes capítulos abordaremos estos elementos definitorios de manera más detallada, construyendo una metodología específica que nos permita abordar la tarea de evaluar el impacto de las políticas de la Junta de Andalucía de una manera lo más asequible posible y que, a la vez, permita unas conclusiones solventes.

Articulación de los elementos definitorios de la Evaluación de Impacto

Los elementos fundamentales que definen la evaluación de impacto que se propone en esta guía se llevarían a la práctica de la siguiente manera:

- **Situación de partida**: Definido por pocos indicadores de contexto, que constituirán la Línea Base (ver capítulo 3) de manera que, en principio, no se tendrán en cuenta efectos sobre elementos del contexto no previstos por la actuación.
- **Temporalidad**: Los efectos se medirán después de la ejecución del programa (Evaluación ex-post) y en el momento en que estos se hayan podido producir (maduración del efecto)
- **Causalidad**: Se considerará que el efecto puede ser causa del programa cuando este tenga una lógica que explique de forma plausible el recorrido hasta llegar a él y cuando, mediante el análisis de su contrafactual, se puede llegar a la conclusión de que el programa ha contribuido a generar el efecto estudiado.

Pero antes de ello, veamos en qué casos nos podríamos decidir por hacer este tipo de evaluaciones, evitando

realizar un sobreesfuerzo con una recompensa limitada o nula. No todas las actuaciones permitirán que se pueda evaluar su impacto, o la cantidad de recursos que demandarían no justificaría semejante esfuerzo.

¿CUÁNDO MEDIR EL IMPACTO?

Como se ha visto en este capítulo, y profundizaremos en los siguientes, la evaluación de impacto no es una evaluación sencilla y, en ocasiones, puede ser frustrante por la falta de consistencia de sus resultados. Aun así, podemos decir que es el tipo de evaluación que más información nos puede proporcionar acerca de los mecanismos de funcionamiento de la acción pública, y de su capacidad de promover cambios en la población o el contexto al que se dirige.

A pesar de esta enorme utilidad y capacidad pedagógica, no siempre será posible abordar este tipo de evaluaciones. Veamos algunas de las cuestiones que nos ayudarán a decidir cuándo afrontar la evaluación de impacto de un programa.

Criterios relacionados con el diseño y la implementación del programa

El primer conjunto de factores que nos van a servir para valorar si emprender o no la evaluación de impacto, está directamente ligado al diseño y la implementación del programa. Un plan con un diseño o con una puesta en marcha deficiente, difícilmente permitirá realizar una evaluación solvente. Además de la influencia de estos propios factores en los resultados finales y en los efectos subsiguientes. Por lo tanto, los principales elementos a valorar son:

Diseño del programa: Un programa con una *lógica de actuación* identificada y plausible, con objetivos definidos, una estructura coherente y un diagnóstico adecuado, nos permitirá afrontar este tipo de evaluaciones con solvencia. Si no cuenta con estas características, habrá que valorar la posibilidad de si puede reconstruirse con cierta facilidad.

Estabilidad de la implementación: Cuanto más estable haya sido el programa durante su ejecución, en el sentido de *no haber sufrido cambios constantes en las actuaciones o en los objetivos*, más oportunidades tendremos de que las conclusiones de la evaluación sean veraces y se correspondan con la actuación llevada a cabo. Esto implica que se ha llevado a cabo conforme a lo previsto o, en su caso, que las desviaciones que ha sufrido están bien documentadas, definidas y justificadas.

Una excepción a este aspecto podría ser en el caso de *proyectos piloto* que están pensados para valorar su efectividad a pequeña escala antes de aplicarlos de manera generalizada. Estos proyectos, por su propia naturaleza y su aspecto pedagógico, pueden estar sujetos a cambios y reformulaciones con cierta frecuencia y será necesario valorar su impacto para poder extraer conclusiones sobre su utilidad y la posibilidad, o no, de que sean generalizados.

Calendario: El tiempo de ejecución es un aspecto clave para abordar la evaluación en el momento en que los efectos se hayan podido producir. Habrá que valorar, a partir del propio programa, *si el tiempo transcurrido es el adecuado*, es decir, no ha pasado demasiado tiempo que

pudiera haber diluido los efectos, ni es demasiado temprano como para que aún no sean observables en todo su desarrollo. Este es un elemento para estudiar detenidamente y poder determinar el momento de la evaluación o, incluso, si ya no es posible hacerla.

Criterios relacionados con el análisis de los efectos y el contrafactual

Este segundo grupo de factores atienden a la factibilidad de la propia evaluación, de la disponibilidad de métodos y datos que nos permitan abordarla. Entre ellos destacamos los siguientes:

Definición de la situación de partida: Un programa con un buen diagnóstico, con un *conjunto de indicadores* de contexto adecuado y que ofrezcan información veraz y fiable, es el "*santo grial*" de la evaluación de impacto. Si no conocemos con detalle la situación de partida, ni tenemos manera de valorar sus cambios, será imposible hacer este tipo de evaluaciones. Al igual que con los programas con diseños deficientes, se tendrá que valorar la posibilidad de reproducir *a posteriori* estos elementos de manera solvente.

Existencia del contrafactual adecuado: Este es un aspecto espinoso ya que, salvo que lo hayamos previsto de manera explícita, lo normal es que no encontraremos poblaciones iguales para comparar. Aunque existen otras formas para analizar el contrafactual que veremos más adelante, este es un aspecto central a la hora de decidir si afrontar estas evaluaciones. Optemos por el tipo de contrafactual que optemos, su análisis requerirá tomar la decisión desde la óptica del *coste-beneficio* que brindará la evaluación en estos casos, dado que suele ser una tarea costosa.

Un caso extremo serían los programas universales, que por estar dirigidos a toda la población difícilmente permitirán una comparación robusta (si no hay elementos de comparación porque toda la población es beneficiaria difícilmente encontraremos un contrafactual adecuado).

Criterios relacionados con la propia política pública

Entre estos elementos de decisión existen dos cuestiones fundamentales, la importancia y estabilidad de la propia política y la necesidad de aprendizaje sobre sus efectos.

Importancia de la política pública: Es casi de Perogrullo, pero la importancia (urgencia, gravedad, etc.) que tenga o se le otorgue a una actuación será el primer elemento a considerar. Sin embargo, no es tan evidente dicha importancia de cara a la evaluación de impacto. Por ejemplo, en el caso de planes o políticas coyunturales, sin estabilidad ni continuidad en el tiempo, puede no ser necesario el esfuerzo hacer estas evaluaciones dado que no volverán a repetirse. El caso contrario, sería para políticas con grandes implicaciones a nivel poblacional, actualidad y trascendencia de los problemas que abordan, para las que sería casi obligada una evaluación de impacto.

Aprendizaje: La necesidad de conocer cómo actúan y qué efectos tienen las acciones públicas, que se han puesto en marcha por parte de la Administración, es otro factor de gran importancia. En estos casos, el esfuerzo puede estar sobradamente justificado, sobre todo, si se evita seguir reproduciendo actuaciones cuyos efectos son más que discutibles.

Tanto la importancia de la política como el aprendizaje son variables que se pueden considerar simultáneamente a la hora de decidir si abordar o no la evaluación de impacto de un programa.

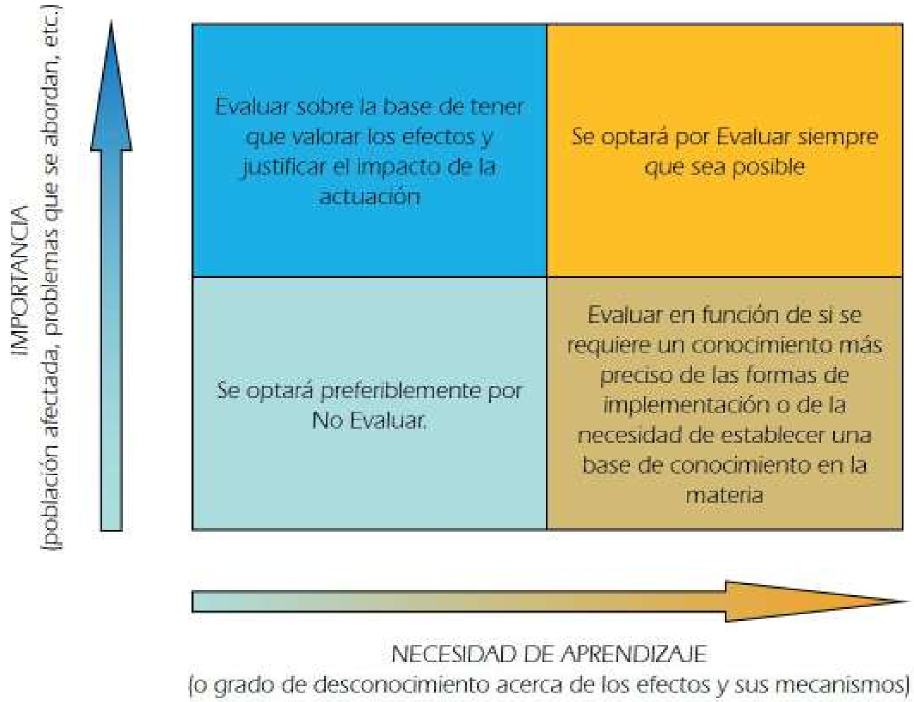


Figura 10: Matriz de decisión según la importancia y las necesidades de aprendizaje

02

CAPÍTULO 2: EN BUSCA DE LA CAUSALIDAD: EL CONTEXTO COMO PUNTO DE PARTIDA Y LA LÓGICA DEL PROGRAMA

Como hemos visto en el capítulo anterior, la clave de la evaluación de impacto es establecer el grado de contribución de una actuación a la consecución de los objetivos que se ha planteado conseguir. Es decir, valorar en qué manera la actuación ha sido la causa de los cambios que se han producido tras su implementación.

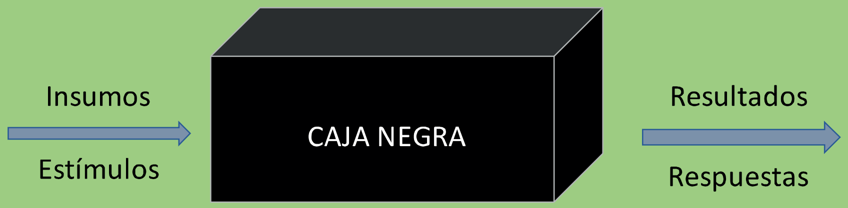
En este capítulo abordaremos con más detenimiento los elementos fundamentales que nos permitirán establecer y valorar dicha relación causa-efecto entre la política pública y los cambios que se ha producido en la población o en el contexto al que iba dirigida.

Recordando lo expuesto anteriormente, estos elementos son dos: La causalidad y el contrafactual. En el primer caso, determinaremos la causalidad mediante el análisis de la "lógica del programa". En el segundo, valoraremos los cambios mediante el establecimiento de la "Línea Base" y su posterior análisis.

Cuando vamos a realizar una evaluación de impacto, nos podemos encontrar con un programa ya implementado o con un plan aún en fase de diseño, en el que podamos valorar tanto su lógica como su contexto, e introducir los cambios que se estimen oportunos. Lo que se expone en los siguientes epígrafes es válido tanto para un caso como para el otro, la diferencia estriba en la dificultad que entraña reconstruir a posteriori la relación causa-efecto de las actuaciones, o que no se hayan contemplado previamente los indicadores de contexto (o de impacto) que se estimen necesarios para valorar los efectos del programa y, por lo tanto, carezcamos de los datos necesarios y tengamos que obtenerlos con posterioridad a la actuación. Es en estos casos, en los que es más importante tener clara la utilidad de abordar este tipo de evaluaciones y, sobre todo, estar seguros de que podremos disponer de los datos necesarios para poder realizarla (indicadores suficientes, datos que alcancen el momento previo a la implementación del programa, etc.)

EL MISTERIO DE LA CAJA NEGRA

En algunas disciplinas, como el análisis de sistemas, existe un concepto que se conoce como la Caja Negra. Según éste, lo importante del sistema no es su funcionamiento sino sus resultados, por lo que se analiza la relación entre los insumos que se introducen en él y los productos que se obtienen.



Un plan estratégico tipo "caja negra", sería aquel en el que nos guiamos por sus resultados sin importarnos cómo se han conseguido. En el caso de que no se haya podido establecer una programación lógica que, sin embargo, parece haber causado los efectos esperados, debemos plantearnos un análisis exhaustivo para poder desentrañar cómo ha actuado, evitando la tentación de seguir aplicándolo en acciones futuras con la confianza de que, sin saber por qué, funciona.

ANÁLISIS CAUSAL: LA LÓGICA DE LA ACTUACIÓN

La lógica de toda actuación parte de la base de la existencia de relaciones causa-efecto del estilo; si se lleva a cabo A se debería conseguir B y B colabora a conseguir el objetivo marcado. Aunque no existen **procesos lógicos** que garanticen los resultados al 100%, sí debemos asegurarnos de que son **consistentes**, apoyándonos en experiencias previas, estudios e investigaciones, es decir, **basados en evidencias**.

Aunque hay varios métodos que nos permiten formalizar y representar las relaciones causa y efecto, los dos principales que se proponen para el diseño de políticas públicas impulsado por el IAAP, son el "**Enfoque del Marco Lógico**"y la"**Teoría del Programa**".

Estos métodos se exponen como herramientas que nos pueden facilitar la labor de valorar o reconstruir la lógica de las actuaciones que subyace en el programa que se va a evaluar. Se han seleccionado pensando en la labor "forense" a la que debe enfrentarse el equipo evaluador sobre un programa ya terminado, y porque permiten de manera sencilla reconstruir la lógica del programa de manera directa a partir de sus componentes (objetivos, resultados,

medidas, etc.) De esta forma, tomando estos elementos y relacionándolos como se muestra en estos métodos, podremos concluir si realmente es lógica o no la estructura del programa que vamos a evaluar, y si podemos afirmar que realmente se ha programado sobre evidencias y relaciones causa-efecto solventes.

Si deseamos profundizar, no solamente en la lógica que subyace en las actuaciones, sino que queremos analizar el motor de cambio sobre el que se apoyan, se recomienda recurrir a la **Teoría del Cambio** de la que se ofrece un ejemplo en la sección "Para saber más".

El Enfoque del Marco Lógico

El enfoque del Marco Lógico (EML), es una metodología relacionada con el diseño de programas públicos, definida inicialmente en el ámbito de la cooperación al desarrollo[8]. El EML parte de la existencia de dos "lógicas" dentro de un programa. Por un lado, la lógica vertical, que muestra la relación causal entre los distintos niveles de programación de la actuación; objetivos, medidas, actuaciones, etc., por otro lado, aborda la lógica horizontal, que muestra la relación entre los elementos que constituyen cada uno de estos niveles; medidas, recursos, condicionantes, etc.

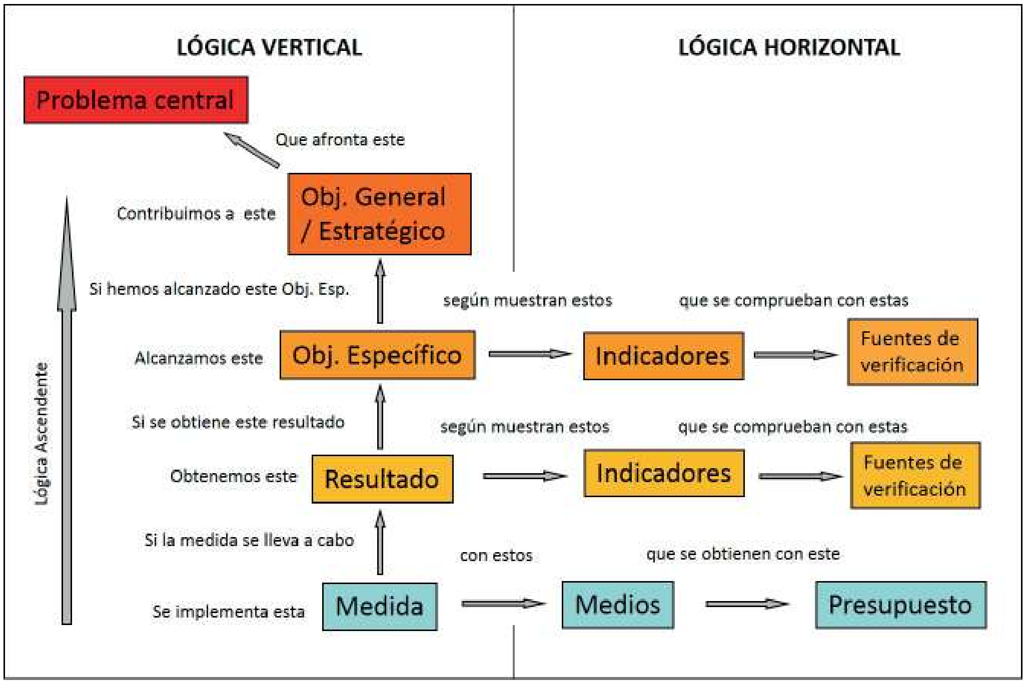


Figura 11: Lógicas vertical y horizontal en la metodología del Enfoque del Marco Lógico

8. *COMISIÓN EUROPEA – EuropeAid, 2001*

En el modelo de evaluación de impacto que se propone en esta guía, es el análisis de la lógica vertical y, por lo tanto, del enlace de los distintos niveles programáticos, la que nos permite comprobar si el planteamiento del programa tiene una cadena de relaciones causales solvente. Mientras que la lógica horizontal, que enlaza todos los elementos de un mismo nivel de programación, nos permitirá, más adelante y sobre la base de los resultados de la evaluación de impacto, estudiar la adecuación de los medios y de los elementos intermedios (productos, resultados, etc.) en cada nivel de la intervención. Por ejemplo, en el caso de que un efecto no haya alcanzado el nivel esperado (posible impacto menor del previsto) se puede deber a que los recursos empleados no han sido suficientes.

La Teoría del Programa

Al igual que en el EML, la Teoría del Programa[9] parte de la base de que de una secuencia lógica se derivan unos determinados productos y resultados que contribuyen a la consecución de los objetivos propuestos en un plan estratégico.

ENFOQUE SISTÉMICO DE LA TEORÍA DEL PROGRAMA

"El enfoque sistémico entiende el programa o servicio como un conjunto de elementos articulados, e interdependientes entre sí, que conforman un sistema. Las relaciones medios-fines recorren el sistema desde su inicio hasta los resultados finales implicando a todos los elementos de la intervención. La variación en uno de los elementos afecta al resto de los componentes."

(Ligero Lasa, 2011:30)

Por lo tanto, en este caso, se da una diferencia sutil pero importante respecto al enfoque del Marco Lógico. Lo que en el primero se separaba en dos lógicas diferenciadas (vertical y horizontal) ahora, en el marco de la Teoría del Programa, se trata en una cadena única con elementos que se alimentan unos a otros hasta llegar al efecto final. Esta cadena puede representarse de diversas maneras, pero se ha optado por el ejemplo que plantea Ligero Lasa de una evaluación de la Unión Europea, por entender que sería la que más facilitaría la labor de reconstruir la lógica del programa.

9. Ligero Lasa, 2011

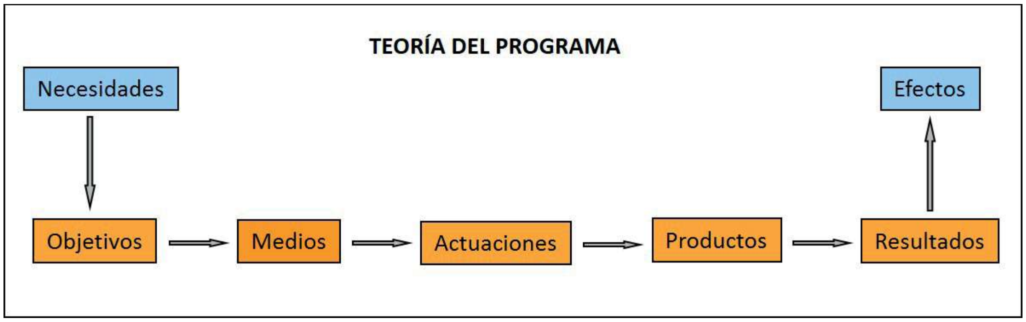


Figura 12: Ejemplo de modelo lógico sistémico elaborado a partir de la Evaluación del Programa de Gastos de la UE de 1997 (Ligero Lasa, 2011:31)

El Análisis Causal

Las metodologías que se han descrito en los epígrafes anteriores nos van a permitir valorar la causalidad del plan que vamos a evaluar. Si bien el Marco Lógico permite un análisis algo más diferenciado de los elementos del programa y la Teoría del Programa sistémica una concatenación más sintética de los elementos, en nuestra opinión, no supone que una sea mejor que otra.

Como ya se ha señalado, es más probable que la evaluación de impacto se realice sobre un programa ya ejecutado,en el cual podemos encontrarnos que no existe una lógica de manera explícita. En estos casos será necesario reconstruirla a partir de los diferentes elementos recogidos en el plan. Para ello se establecerá la cadena de relaciones entre los componentes del plan estratégico siguiendo el modelo lógico que mejor se adapte a nuestras necesidades.

Por otro lado, también podemos encontrarnos con un plan estratégico con una lógica explícita o en el que ya se haya previsto la relación causal de cara a una futura evaluación. Este caso ideal, permitirá poder emplearla directamente sin tener que recurrir a los métodos expuestos anteriormente.

Preguntas para el análisis causal

Para el análisis causal se recomienda recurrir a preguntas sencillas que nos permitan orientar el análisis como, por ejemplo:

Sobre el diagnóstico del contexto (problemas y necesidades):

¿Qué provoca que el problema ocurra?

¿Cuál es la causa de que...?

¿Qué implicaciones tiene ...?

¿Por qué se da la situación...?

Sobre los elementos del programa (objetivos, resultados y medidas):

¿Qué necesitamos para alcanzar ...?

¿Cómo se soluciona ...?

¿Qué nos permite lograr este...?

¿A qué colabora/contribuye...?

¿Qué consecuencias tiene...?

Una vez que tenemos identificada la lógica del programa y ordenada la secuencia de relaciones causa-efecto, procederemos a determinar su solven-

cia. Es decir, intentaremos establecer si esa cadena es coherente y si existen evidencias que soporten la lógica de la actuación permitiendo afirmar que, *a priori*, cada actuación puede realmente colaborar en la consecución del objetivo que persigue o que cada producto podrá permitir la obtención del resultado esperado.

Para ello, recurriremos a realizar preguntas sobre la lógica de cada eslabón, las cuales deberán responderse sobre la base de las evidencias que hubiera en la materia.

Siguiendo con el ejemplo sobre apoyo a la renta, imaginemos que el programa para llevarlo a cabo tiene la siguiente lógica:

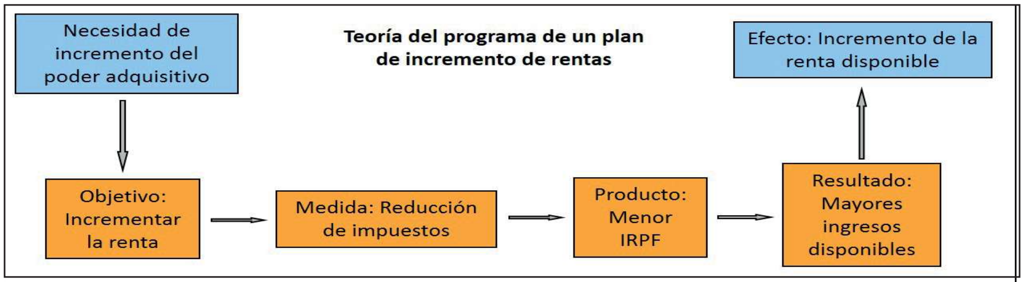


Figura 13: Ejemplo simplificado de modelo lógico para un plan de incremento de rentas

En este caso algunas de las preguntas que nos tendríamos que hacer son:

- ¿El objetivo es pertinente con las necesidades, problema o retos detectados?
- ¿La medida adoptada ha podido lograr el objetivo establecido?
- ¿Un menor IRPF contribuye a un incremento de los ingresos? ¿De qué forma?

Para algunas de estas preguntas no necesitaremos recurrir a un análisis exhaustivo, pero puede que otras no sean tan evidentes y requieran un estudio pormenorizado de las evidencias disponibles (experiencias previas, investigaciones, etc.) para poder alcanzar una respuesta.

DETECTANDO FALLAS DEL PROGRAMA

Dentro de este análisis sistemático de la causalidad o de la lógica del programa es fundamental para identificar posibles lagunas y fallas del programa que condicionarán o podrían impedir la evaluación de impacto, entre ellas la existencia de **Objetivos y Medidas**:

- **Inalcanzables**: Demasiado ambiciosos o fuera de la competencia del programa
- **Incoherentes**: No son pertinentes a lo que se desea conseguir y, por lo tanto, no responden a una relación causal o a la lógica del programa.
- **Huérfanos**: Sin líneas estratégicas, medidas o actuaciones.
- **Ocultos**: Objetivos y medidas implícitos, no explicitados en el programa

LA LÍNEA BASE: DESCRIBIENDO EL CONTEXTO

Una vez establecida y comprobada la lógica del programa y, por lo tanto, la cadena de resultados que se podría obtener de la aplicación del plan estratégico que se evalúa, podemos pasar a analizar el conjunto de indicadores que nos permitirán valorar los efectos del plan sobre la población a la que se dirige. Estos indicadores constituyen la Línea Base del programa que, como ya se señaló en la aclaración metodológica, la definimos de la siguiente manera:

> *La **Línea Base** de un programa es un conjunto limitado de indicadores de contexto que permite valorar los cambios fruto promovidos por la actuación en la población o el contexto en la que se aplica, debiendo poder describir su situación inicial y final en relación con los objetivos perseguidos por el programa.*

Si los indicadores que constituyen la Línea Base ya están definidos en el programa procederemos a confirmar su pertinencia, si no es así, deberemos determinar cuáles son los más indicados. En este apartado partiremos de la premisa de que debemos establecer la Línea Base, aunque los pasos que se muestran también son útiles para confirmar si, en caso de estar ya definida, es pertinente con los objetivos del programa.

Para ello, tomaremos como referencia los objetivos estratégicos del programa, ya que son ellos los que precisan el efecto que se quiere causar con la implementación de las medidas recogidas en el mismo. De esta manera, y tras comprobar en el paso anterior que tienen coherencia y son pertinentes al diagnóstico, a partir de su análisis pormenorizado podremos proponer una batería de indicadores de contexto con los que valorar los efectos del plan. Veamos cómo abordar este análisis de los objetivos.

Diseccionando los Objetivos Estratégicos

El análisis de los objetivos, para determinar los efectos que se pretenden causar sobre el contexto diagnosticado, parte de determinar las dimensiones de que se componen.

DIMENSIONES DE UN OBJETIVO

Entendemos por dimensiones de un objetivo aquellos elementos clave y diferenciables que se pretenden conseguir con él. En la práctica se corresponden con los elementos del diagnóstico detectados (necesidades o problemas) que, o bien se podrían abordar de manera individualizada, o bien son la suma de varios aspectos individualizables.

En función de las dimensiones que tenga un objetivo, nos podemos encontrar con:

Objetivos unidimensionales: aquellos que sólo acometen un problema o necesidad detectada.

Objetivos multidimensionales: aquellos que abordan dos o más problemas o necesidades detectadas o son de naturaleza compleja al reunir a su vez diferentes aspectos clave que deben analizarse separadamente.

Para analizar este último tipo de objetivos los dividiremos en función de dichas dimensiones para tratarlas de manera separada, dado que representan efectos diferentes. De esta manera mediante el análisis separado de cada dimensión, se determinarán qué indicadores de contexto serán los necesarios para valorar los cambios en cada problema o necesidad que abarca el objetivo estratégico.

Tomando como base el ejemplo del programa de apoyo a la renta, veamos, de una manera simplificada cada caso, un objetivo unidimensional y otro bidimensional.

EJEMPLO DE OBJETIVO UNIDIMENSIONAL

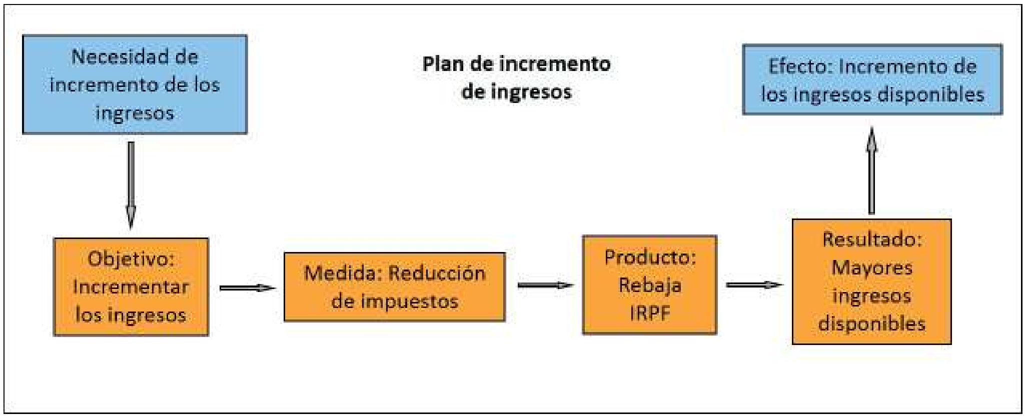


Figura 14: Ejemplo simplificado de plan de incremento de ingresos

En este caso, en el programa se ha marcado el objetivo de incrementar los ingresos, el cual tendría una única dimensión, los ingresos de las familias. Por lo tanto, para ver el efecto del programa nos bastaría con mirar a los cambios en los ingresos de las unidades familiares.

EJEMPLO DE OBJETIVO MULTIDIMENSIONAL

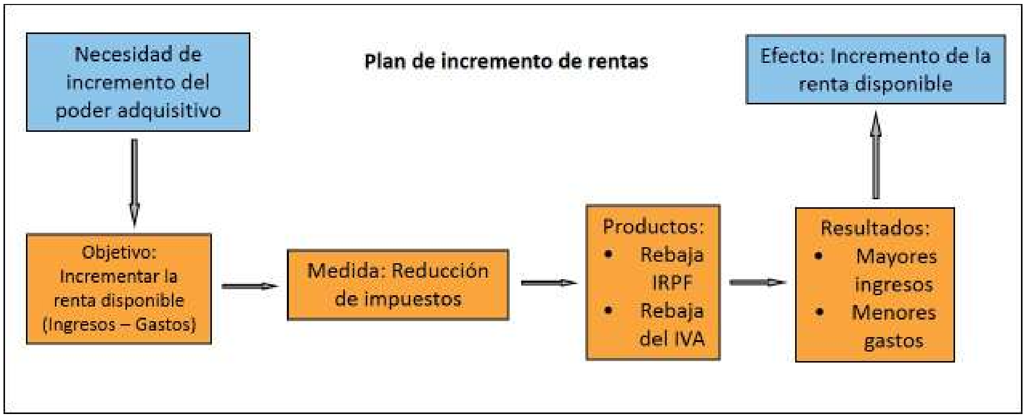


Figura 15: Ejemplo simplificado de plan de incremento de rentas

En este segundo caso, el objetivo sería el incremento de la renta disponible, por lo que tendría dos dimensiones o necesidades que afrontar al mismo tiempo; por un lado, los ingresos y, por otro, los gastos. Por lo tanto, a la hora de analizar los efectos de este plan debemos atender los efectos en ambos elementos clave.

Selección de indicadores de la Línea Base

Una vez que tenemos identificadas las dimensiones de los objetivos estratégicos de nuestro plan, pasamos a analizar qué indicadores nos pueden dar información relevante sobre los cambios que se produzcan en dichas dimensiones.

Figura 16: Tipos de indicadores según el elemento del plan sobre el que suministran información

Los indicadores de Impacto que se seleccionan para evaluar los objetivos estratégicos son los mismos que los empleados para describir el contexto y realizar el diagnóstico de la situación que se aborda en el plan. Sin embargo, hay que tener en cuenta que no todos los indicadores de contexto servirán para medir el impacto. Esto es, generalmente, porque en el diagnóstico se hace un análisis amplio de la población, sus problemas y necesidades (empleando dichos indicadores de contexto) mientras que los objetivos estratégicos implican una decisión restrictiva de qué problemas o necesidades se van o se pueden atender. Es decir, en circunstancias normales un plan estratégico no se va a plantear como la solución a todo.

INDICADORES DE IMPACTO

En el modelo que se plantea en este manual, entenderemos que los indicadores de impacto son aquellos indicadores de contexto que nos brindan información pertinente sobre los cambios en los elementos que se pretenden afrontar por los objetivos estratégicos del plan o programa

Los indicadores deberán ser válidos, es decir, suministrar información representativa de la situación sobre la que valoramos los efectos, ser **fiables**, es decir, que la información sea coherente en todo momento, suministrando el mínimo posible de valores erróneos, y ser **estables en el tiempo**, es decir, que nos den dicha información tanto en el momento de la evaluación, como otros momentos de la implementación del programa, especialmente antes de la misma.

Este conjunto reducido de indicadores de contexto será lo que constituya nuestra Línea Base. En el ejemplo visto anteriormente, para el objetivo multidimensional sobre incremento de rentas podrían ser dos indicadores estadísticos sobre las cuentas en los hogares.

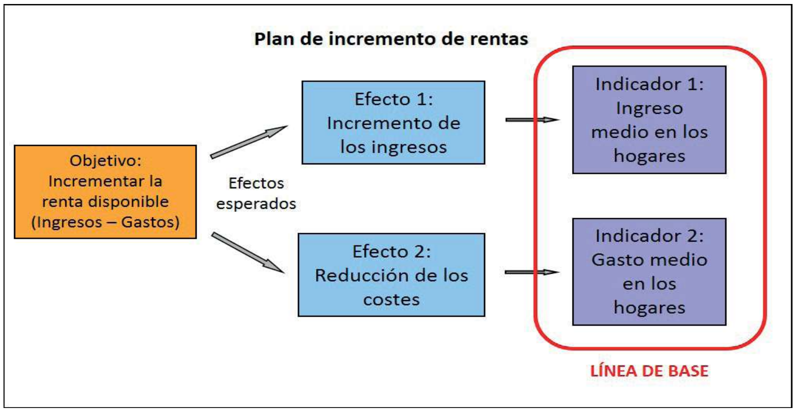


Figura 17: Línea Base simplificada del plan de incremento de rentas

Finalmente, antes de proceder al análisis contrafactual sobre la base de estos indicadores, deberemos establecer los valores de partida de esta Línea Base. En este punto nos podemos encontrar con dos tipos de líneas de base:

Línea Base estática: Es aquel conjunto de indicadores de contexto para el que solamente se establece el valor antes de la ejecución del programa.

Línea Base proyectada: Es aquel conjunto de indicadores de contexto para el que se establece tanto el valor inicial del programa, como los valores teóricos esperados tras su implementación, que se habrán estimado cuando aún no se ha ejecutado el programa, durante la evaluación ex ante.

Aunque el uso principal de este último tipo es la fijación de metas durante la fase de planificación (resultados esperados del plan, más allá de la evolución natural de los indicadores) como se verá más adelante, este segundo tipo de Línea Base podría suponer la plataforma para una evaluación de impacto, ya que permite ser utilizada como contrafactual.

Para ilustrar la Línea Base y sus tipos, tomemos el ejemplo más sencillo del objetivo unidimensional de incremento de ingresos que mostrábamos anteriormente, en el que tenemos unos datos iniciales “estáticos”[10] (azul) y unos datos proyectados antes de la ejecución del plan estratégico (naranja):

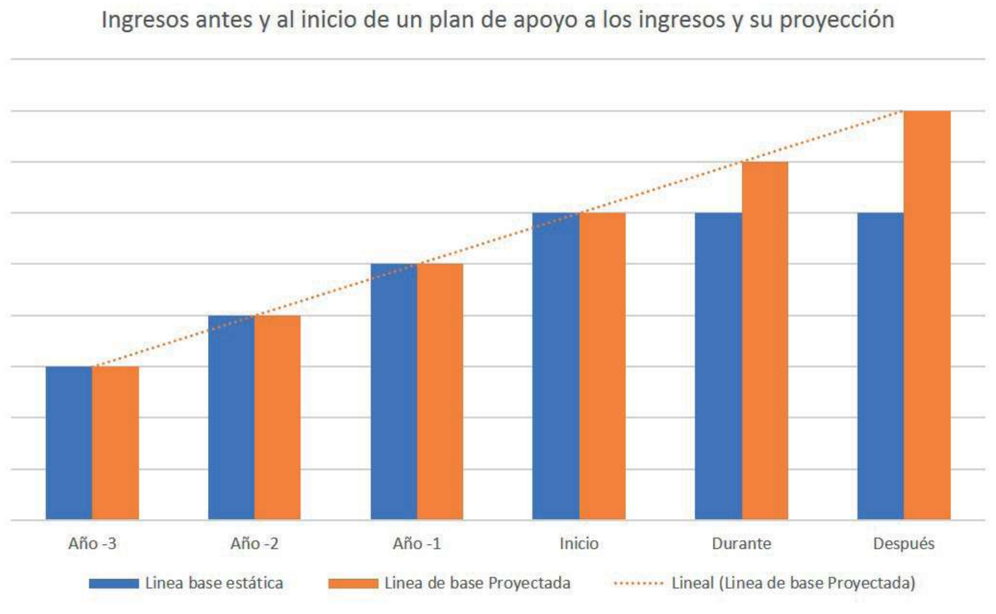


Figura 18: Tipo de líneas de base para un plan de incremento de ingresos

10. Lógicamente los datos se pueden actualizar durante el programa y a su finalización, pero ya se verán influidos por él.

Si no se ha explicitado en la evaluación *ex-ante* del programa, la decisión a posteriori de usar uno u otro tipo de Línea Base[11], vendrá determinada fundamentalmente por la disponibilidad de datos que tengamos, y la existencia de métodos de proyección fiables. En este sentido, debemos tener en cuenta que la disponibilidad de datos para los indicadores debe abarcar un periodo de tiempo suficiente, contado hacia atrás desde el inicio del programa, para poder aplicar los modelos de proyección.

Una vez que tenemos comprobadas la lógica del programa y las relaciones-causales y seleccionada la Línea Base, podemos afrontar el análisis del contrafactual, tal y como se aborda en el siguiente capítulo.

11. Aunque preferentemente buscaremos líneas de base proyectadas, que nos suministran más información, es muy probable que no nos resulte fácil o accesible obtenerlas. Como veremos en el siguiente capítulo, en algunos casos, serán estas proyecciones las que nos servirán como contrafactual.

03

CAPÍTULO 3: MIDIENDO EL IMPACTO: DEFINIR EL CONTRAFACTUAL

COMPLEJIDAD Y COMPROMISO

Como veremos a lo largo de este capítulo, nos adentramos en el "hueso" de la evaluación de impacto, la zona donde las herramientas disponibles se vuelven complejas y donde no se encuentran soluciones que cubran toda la casuística que puede envolver este tipo de evaluaciones.

Esta complejidad lleva a tomar una elección basada en el compromiso, en el equilibrio entre los datos, los medios y el tiempo disponibles y la validez y fiabilidad de los resultados de la evaluación.

Una vez que se ha definido de forma clara la lógica de la actuación a evaluar y se conoce qué se quiere medir, se deberá pasar a la parte que, sin duda, es más compleja en el proceso de evaluación de impacto: establecer cómo medir los efectos de manera que se pueda diferenciar entre los que han sido fruto del programa y los que no.[12]

Como ya se avanzó en el capítulo anterior, la piedra angular de la evaluación de impacto, y el elemento que la diferencia del resto de evaluaciones es la medida de los efectos del "no-programa", el conocido como contrafactual. Durante este capítulo vamos a desgranar algunas de las metodologías que consideramos más extendidas y adecuadas para conseguir la información de referencia necesaria para poder valorar los efectos de un plan estratégico.

12. En las metodologías que se exponen en este capítulo se ha intentado mantener separados los conceptos de "población" y "contexto", ya que algunas de las que se muestran son fundamentalmente aplicables a uno de ellos. En estos casos la mención de población o de contexto no es intercambiable del mismo modo que en el resto del texto.

EL CONTRAFACTUAL O CÓMO VALORAR LOS EFECTOS DEL NO PROGRAMA

El contrafactual no es ni más, ni menos, que la referencia de la Línea Base que explicamos en el capítulo anterior, el kilómetro cero sobre el que comparar los progresos del mismo set de indicadores, en el contexto o en la población beneficiaria del plan estratégico a evaluar. Es decir, es el conjunto de efectos, entendidos como variaciones de los indicadores de la Línea Base, no debidos al programa, que nos permite valorar los efectos a los que dicho plan ha contribuido. En este sentido, el contrafactual es el "grupo de control" contra el que comparar los efectos de una acción pública.

Entre los contrafactuales a los que podemos recurrir para poder hacer esta comparación tenemos dos grandes opciones:

Grupos de Control formados por la población que no ha participado en el programa. En este tipo de contrafactual se pueden obtener los valores de los efectos del "no programa".

Contrafactuales reflexivos en los que las referencias con las que se comparan los efectos del programa provienen de una estimación o de estudios específicos, pero no de los valores obtenidos de una población concreta.

Los primeros son los contrafactuales más robustos y que mejores datos nos pueden ofrecer, sin embargo, conllevan algunas limitaciones, una intensidad de medios y una dificultad metodológica, que hacen que no siempre sean opciones viables para la evaluación de un programa. En todo caso, los tipos de contrafactual que analizaremos en este capítulo son los siguientes:

TIPOS DE CONTRAFACTUAL	
GRUPOS DE CONTROL	REFLEXIVOS
EXPERIMENTALES GRUPOS RECONSTRUIDOS DOBLES DIFERENCIAS	SERIES TEMPORALES ESTUDIOS DE CASO

Validez de los resultados

Una de las cuestiones clave para la elección del contrafactual es la validez de las conclusiones que obtengamos de su análisis, entendida como la capacidad que tiene de darnos información pertinente de los efectos del "no-programa" para obtener conclusiones reales sobre el propio programa. En este sentido debemos distinguir dos tipos de validez:

Validez interna: Hace referencia al contrafactual en sí mismo, a que realmente represente los efectos de

no haber sido población directamente beneficiaria de la política pública que se está evaluando.

Esta validez está condicionada por muchos elementos, desde los factores del entorno (económicos, sociales, educativos, etc.) hasta los relacionados con cuestiones estrictamente metodológicas, como el efecto de la observación sobre la propia población observada[13].

Validez externa: Hace referencia a la capacidad de extrapolar los resultados del análisis al resto de población susceptible de ser objeto de la política pública, es decir, la capacidad de generalizar los resultados obtenidos.

Para asegurar en la mayor medida posible la validez de los resultados de la evaluación, la selección aleatoria es la forma ideal de obtener tanto el grupo de control (contrafactual) como el propio grupo de tratamiento (muestra de personas beneficiarias del programa). El hecho de que las unidades o individuos sobre los que medir los efectos, tanto del programa como del "no programa", se escojan de manera aleatoria, minimiza los riesgos de que se den factores que pongan en peligro la validez de la evaluación, ya que este tipo de selección tiende a anular los errores y sesgos asociados a la selección.

CONTRAFACTUALES CON GRUPOS DE CONTROL

Precisamente los contrafactuales que pueden aportar una mayor validez a las conclusiones de la evaluación de impacto, son aquellos que ofrecen información directa de los efectos de no participar en el programa,a través del estudio de un grupo poblacional que realmente no ha participado en el mismo. Este conjunto de individuos constituye el "grupo de control", frente al "grupo de tratamiento", que incluye a quienes se beneficiaron del programa que se evalúa.

A continuación, vamos a analizar varias metodologías que emplean este tipo de grupos de control para realizar las comparaciones entre los efectos de participar y de no participar en un programa determinado.

Experimentales o aleatorios

En estos casos, el grupo de control, e idealmente el grupo de tratamiento, se selecciona como en un experimento, de manera puramente aleatoria. Esto significa que dentro de la población potencialmente beneficiaria se seleccionan al azar tanto las personas participantes como las no participantes en el programa.

Esta selección aleatoria da una gran robustez al análisis, pero plantea numerosos problemas y limitaciones en el momento de evaluar el impacto, especialmente en programas con ámbitos y poblaciones objetivo amplios. El simple hecho de la selección aleatoria puede suponer un problema a la hora de seleccionar las personas beneficiarias de una acción pública, tanto legal (hay normas que exigen una selección en base a criterios específicos y baremos) como ética y política (puede ser difícilmente entendible que un programa decida la participación de una persona frente a otra por simple azar).

Conocido como Efecto Hawthorne, es el fenómeno por el cual las personas sometidas a un experimento cambian inconscientemente su comportamiento simplemente por participar en él.

Sin embargo, esta metodología es muy recomendable para la evaluación de acciones innovadoras mediante proyectos piloto, que permitan experimentar las intervenciones a pequeña escala antes de decidir su posible extensión a una población más amplia. En todo caso y dado que esta guía va dirigida a la evaluación de impacto de planes estratégicos que, por definición, suelen tener un ámbito más general, no recomendaremos este tipo de contrafactuales o diseños experimentales para su evaluación de impacto.

Grupos reconstruidos: Emparejamiento o Matching

Muchas veces nos encontraremos con la necesidad de evaluar ex-post un programa que no previó un grupo de control específico, o un programa que, como hemos señalado anteriormente, no permite la selección de este tipo de grupos en la fase *ex-ante* del mismo. En estos casos podemos recurrir a reconstruir el grupo de control a posteriori sobre la base de replicar una muestra de personas beneficiarias entre la población que, sin haber participado en el programa, podrían haberlo hecho.

El principal problema de reconstruir el grupo de control una vez terminada la fase de implementación de una acción pública, es encontrar un conjunto de individuos equivalente al grupo de personas beneficiarias del programa. Esta equivalencia es fundamental, ya que si hay factores diferenciales la validez del resultado de la comparación será limitada, especialmente si se sabe que dichos factores tienen trascendencia en la magnitud del efecto que se evalúa. Por ejemplo, factores como la edad o la formación previa, pueden condicionar los resultados de un programa de inserción laboral.

Para conseguir esta equivalencia, la selección del grupo de control se hace

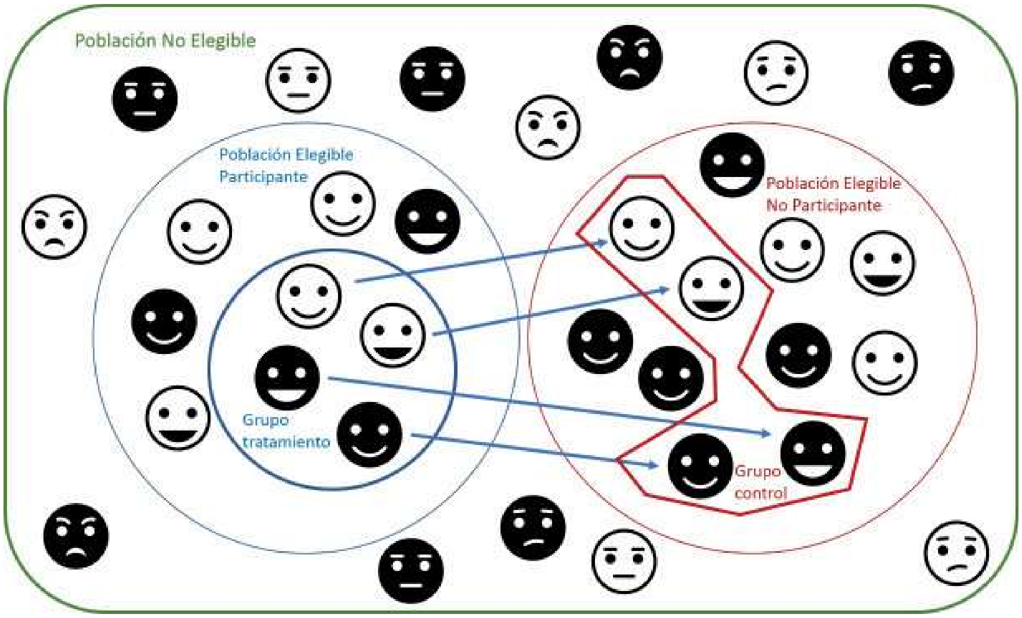


Figura 19: Concepto de la selección de un grupo de control por emparejamiento

sobre la base de dos criterios fundamentales. Primero, que la probabilidad para haber participado en el plan hubiera sido la misma que la de los que realmente participaron y, segundo, que la distribución de factores que pueden influir en los efectos que vamos a medir (edad, sexo, formación, etc.) sea similar a la distribución de estos mismos factores en el grupo de participantes.

Existen diversas técnicas estadísticas para el emparejamiento, que no entraremos a tratar aquí. Las más empleadas con el *Propensity Score Matching* (PSM) que ofrece la probabilidad de participar en función de las características del individuo, y que se puede combinar con la técnica del vecino más próximo (*Nearest-Neighbor Matching*) que permite emparejar individuos con un PSM parecido.

Una vez obtenido el grupo de control se obtendría el conjunto de valores de los indicadores de la Línea Base y se calcularía la diferencia de medias entre ambos grupos, el Grupo de tratamiento y el Grupo de control. Esta diferencia sería el impacto del programa.

Así, por ejemplo, en el caso de un programa para el fomento de la inserción laboral de jóvenes, tras seleccionar aleatoriamente al grupo de participantes del programa, buscaremos un "grupo espejo" de jóvenes que no hayan participado en él, de manera que el impacto será la diferencia del grado de inserción laboral entre ambos.

Figura 20: Medida del impacto mediante un grupo de control por emparejamiento

Finalmente, hay que indicar que este tipo de metodología parte de la base de que las parejas no difieren en factores que puedan influir en el resultado de la medida, es decir que no hemos pasado por alto ningún elemento que pueda condicionar el resultado del grupo de control.

Respecto a la aplicabilidad de este tipo de contrafactual, debemos indicar algo similar a lo dicho con los grupos experi-

mentales, aunque en este caso la complejidad y demanda de recursos que requiere el seguimiento del grupo de control son marcadamente inferiores. Por otro lado, respecto a otras técnicas de grupos de control, esta presenta la enorme ventaja de no necesitar datos del grupo de control anteriores a la implementación del programa. A pesar de esto, siguen teniendo un uso limitado para la evaluación de programas con un ámbito de aplicación amplio y requieren el conocimiento del mayor número de variables posible del grupo de participantes para asegurar su validez.

Dobles Diferencias

Esta es la última de las técnicas de contrafactual que vamos a ver basadas en grupos de control. La metodología de "**diferencias en diferencias**" (término anglosajón) o de "**dobles diferencias**" (DD) se basa en la selección de un grupo de control que hubiera podido acceder al plan evaluado pero que no lo hizo.

Como hemos visto en el caso anterior, hay factores que pueden influir en los resultados, produciendo un sesgo por selección. Para evitarlo o minimizarlo, se toma la variación de los efectos entre "el antes" y "el después" del programa y, esta diferencia "antes-después" del grupo de control se compara con la diferencia "antes-después" del grupo participante. De esta manera, se considera que, independientemente de sus características, los efectos del entorno afectan en el mismo sentido a ambos grupos, que comparten una similitud suficiente por el hecho de ser igualmente elegibles, por lo que si no podemos recurrir a técnicas de emparejamiento o las variables a considerar para aplicarlas

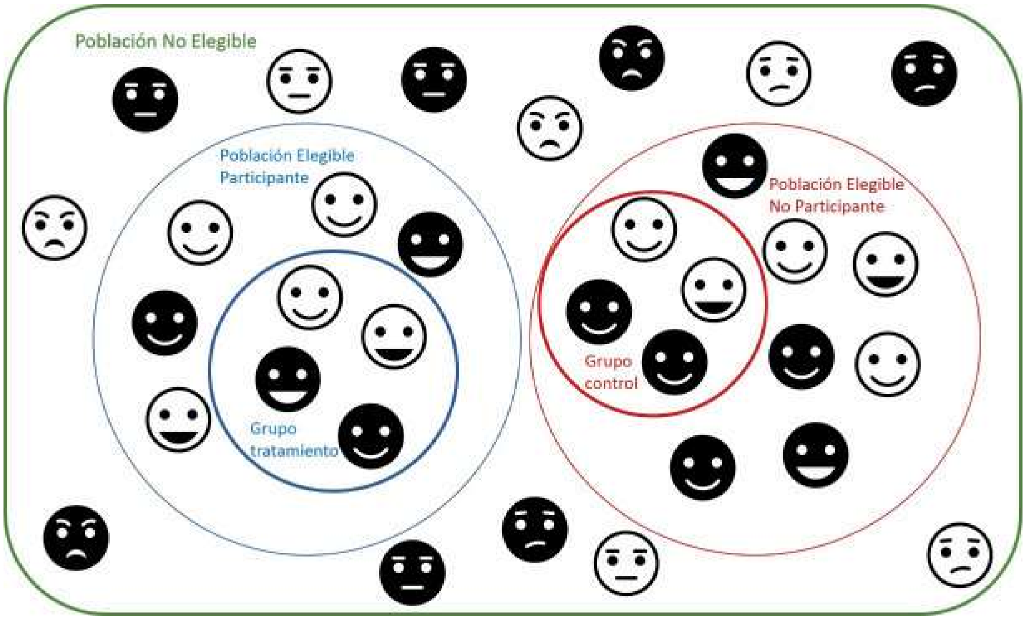


Figura 21: Concepto de la selección de un grupo de control para Dobles Diferencias

son excesivas[14] esta técnica nos puede ofrecer una alternativa aceptable.

Básicamente, la aplicación de esta técnica pasa por una selección del grupo de control de entre las personas beneficiarias potenciales del programa que no han participado en él y, una vez realizada la selección, comparar la evolución del efecto a valorar entre "el antes" y "el después" del programa en ambos grupos (tratamiento y control). Esta comparación se obtiene de restar los cambios de los indicadores antes y después (diferencia de diferencias). Gráficamente, esta doble diferencia supone trasladar la línea de evolución del grupo control al grupo de participantes, para obtener el equivalente a su evolución sino hubieran participado el programa.

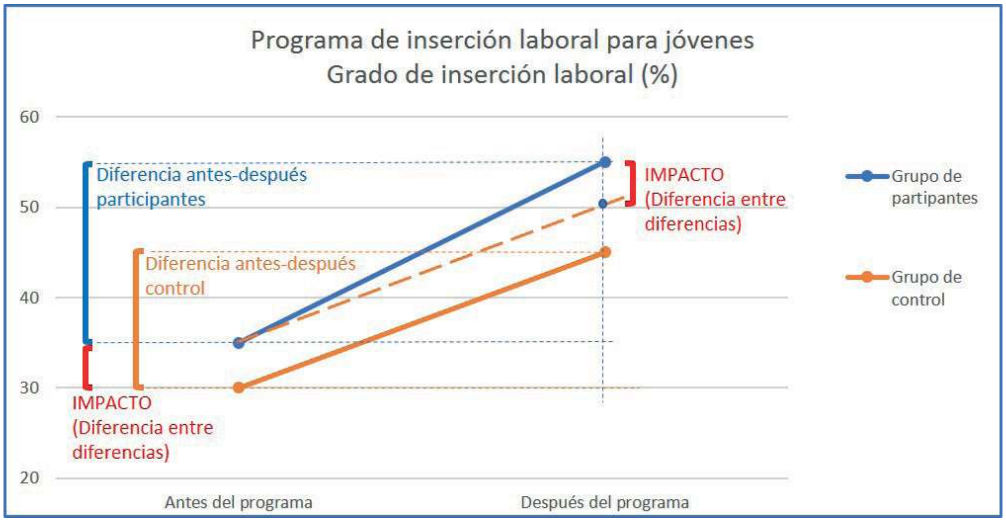


Figura 22: Medida del impacto mediante Dobles Diferencias

Numéricamente el ejemplo anterior sería:

	Programa de inserción laboral para jóvenes Grado de inserción laboral (expresado en %)		
	Antes del programa	Después del programa	Diferencia (Después – Antes)
Grupo de participantes	35	55	20
Grupo de control	30	45	15
Doble Diferencia (Impacto)			**5**

Tabla 1: Ejemplo numérico de medida del impacto mediante Dobles Diferencias

14. Por ejemplo, para evaluar el éxito de un programa de inserción laboral por municipios, las variables a considerar para emparejar cada municipio pueden ser excesivas o, simplemente, no poder encontrar suficientes emparejamientos.

Esta metodología requiere el cumplimiento de dos supuestos, por un lado, que la evolución de los efectos entre ambos grupos sea realmente similar y, por otro lado, que los factores propios de los grupos que puedan influir en estos efectos se mantengan en el tiempo. Es decir, que no haya una diferencia sustancial en elementos clave de la población entre el antes y el después del programa.

En nuestro ejemplo, podría ser el interés por encontrar empleo. Este debe ser el mismo en cada grupo, antes y después del programa.

Por último, hay que señalar que, en los casos en los que sea factible, combinando esta técnica con la de emparejamiento o *matching*, podemos reducir el riesgo de incumplir el primer supuesto (el comportamiento similar de los grupos durante el periodo evaluado).

CONTRAFACTUALES REFLEXIVOS – SIN GRUPOS DE CONTROL

En la labor evaluadora nos encontraremos con no pocas ocasiones en las que, por dificultades técnicas, cuestiones metodológicas, falta de recursos o por simples razones prácticas, no sea posible o recomendable recurrir a grupos de control.

En estos casos disponemos de otras formas de establecer y analizar el contrafactual que, sin la robustez de los casos expuestos anteriormente, nos permiten aproximaciones solventes al impacto de un programa. De entre las posibles opciones metodológicas a las que se puede recurrirse proponen dos: El uso de series temporales para proyectar los efectos del "no-programa" y los estudios de caso.

Series temporales

En el caso de no poder contar con un grupo control de referencia con el que medir los efectos del programa evaluado o en el caso de tener que medirlos sobre un contexto más amplio que un grupo poblacional, podemos recurrir a estimar lo que habría pasado si el programa no hubiera existido a partir de los datos previos a su puesta en marcha. De este modo, se emplean series temporales para predecir el comportamiento de un supuesto grupo de control, o para contrastar la evolución del contexto al margen del programa.

Para conseguir la mayor validez posible de los resultados, se requiere que exista información previa suficiente para que las predicciones sean creíbles y fiables, además de ser necesario que las metodologías de predicción (modelos) estén contrastados y aceptados de manera general por la comunidad científica.

Como se puede observar en la gráfica, la medida del impacto proviene de comparar el efecto en la población beneficiaria o en el contexto de la acción pública, con el que supuestamente se produciría de manera "natural" sin la presencia del programa.

Evidentemente no todos los modelos corresponderán a regresiones lineales como la del ejemplo del programa de inserción laboral. Los modelos podrán variar y debemos hacer una labor investigadora profunda para determinar el que mejor resultado nos pueda ofrecer, según el tipo de efecto que se va a medir y los datos de los que se dispone. Así podemos encontrarnos con diferentes tipos de modelos (geométricos,

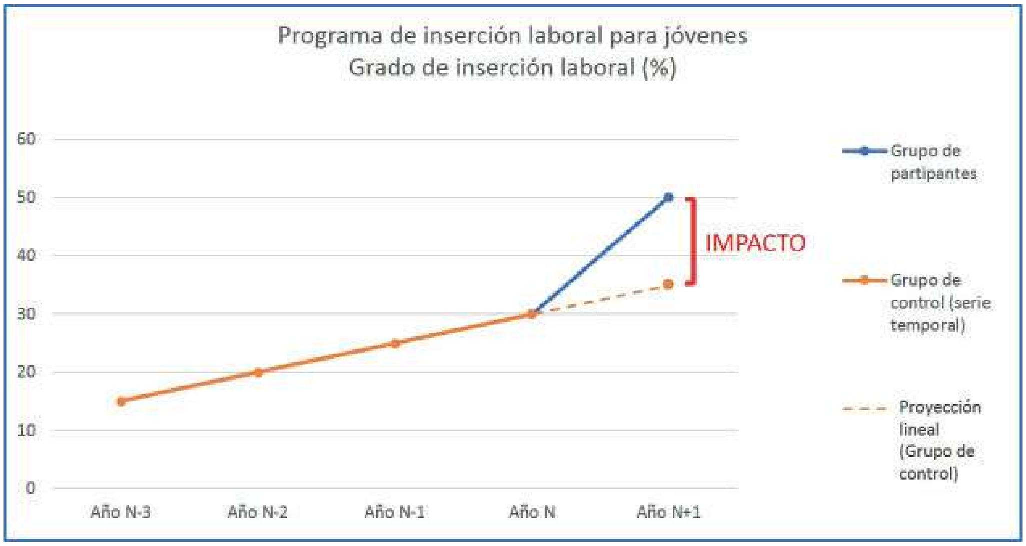


Figura 23: Medida del impacto mediante series temporales

logarítmicos, etc.) y funciones de distribución del efecto que vamos a estudiar. Por ejemplo, para la estimación de variables económicas se pueden utilizar diversos modelos econométricos, para cuya elección deberemos considerar la variable a estimar, el sector sobre el que se aplica, los datos disponibles, etc.

Por lo tanto, el proceso de evaluación parte del análisis de los indicadores de la Línea Base que definen la población o el contexto de partida, comprobar la disponibilidad de datos fiables sobre los mismos, seleccionar los mejores modelos disponibles y proyectar la evolución mediante dichos modelos. El resultado será nuestro equivalente al grupo de control y la diferencia entre éste y los indicadores de la población participante será nuestra medida de impacto.

Como puede observarse, si partimos de un plan con una Línea Base proyectada que se indica en el último apartado del capítulo anterior, tendremos mucho avanzado y, realmente, sólo nos quedará comprobar si el modelo aplicado y la proyección obtenida mantienen su validez.

En cuanto a los pros y contras de este método, hay que indicar que es el menos robusto de los vistos hasta ahora porque parte de la base de que el comportamiento del hipotético grupo de control, se mantiene bastante estable en el tiempo, lo que permite proyectar su situación futura a partir de datos pasados, y que, al igual que el caso de las dobles diferencias, la evolución del grupo de personas beneficiarios o grupo de tratamiento, en el caso de no haber participado en el plan, sería la misma que la proyección realizada.

Teniendo en cuenta lo anterior, se podría recomendar el uso de este contrafactual, cuando no siendo posible construir un grupo de control, se tenga cierta seguridad de que los factores externos que influyen en la evolución temporal de los efectos puedan recogerse en el modelo de proyección.

Finalmente, hay que señalar que, si bien este contrafactual es el más sencillo de los planteados hasta ahora, desde el

punto de vista metodológico se requiere una buena formación estadística para poder abordarlos con solvencia.

Estudios de caso

Aunque realmente este no es un método para construir un contrafactual, no podemos dejar de mencionar esta opción para casos en los que, no solamente no es posible construir un grupo de control, sino que tampoco nos permiten el uso de modelos matemáticos, bien por falta de modelos predictivos solventes, bien por insuficiencia de datos para obtener proyecciones válidas.

Esta situación nos la podemos encontrar al enfrentarnos a la evaluación de impacto de iniciativas muy novedosas, para las que no hay datos previos, y cuya naturaleza no permita establecer grupos de control válidos. O también, al enfrentarnos a planes cuya lógica es insuficiente o su implementación ha sufrido variaciones importantes, rompiendo con las premisas de estabilidad temporal de los factores de los grupos de control.

Aunque estas no son, ni mucho menos, las únicas cuestiones a tener en cuenta, lo cierto es que, ante planes con las limitaciones anteriormente expuestas, la mejor opción son los estudios de caso, es decir, el análisis comparado con otras iniciativas similares, con objetivos comunes o con intervenciones equivalentes.

RAZONES PARA USAR EL ESTUDIO COMPARATIVO DE CASOS

"Cuando se plantean preguntas relativas al «cómo» y el «por qué» de los procesos o los resultados de una intervención.

Cuando se ejecutan una o más intervenciones en varios contextos y hay pocas oportunidades, o ninguna, de manipular o controlar la forma en que se implementan.

Cuando existe la posibilidad de recolectar y analizar datos iterativos durante el marco de tiempo de la intervención.

Cuando se considera que la comprensión del contexto es importante para entender el éxito o el fracaso de la intervención.

Cuando los diseños experimentales o cuasiexperimentales son inviables por razones prácticas o éticas, o para complementar las pruebas de dichos diseños de evaluación."

(Goodrick, D. 2014:02)

Sin entrar en grandes profundidades, el estudio comparativo de casos pasa por una investigación profunda del programa a evaluar: objetivos, ámbito de actuación, población a la que se dirige, teoría del cambio, etc. para poder realizar una búsqueda y revisión de otras iniciativas de similares características.

Previamente a entrar a comparar cada caso, debemos tener claro qué elementos comunes vamos a considerar y qué efectos vamos a valorar, para lo cual nos podemos ayudar de un cuestionario previo al que someter a cada programa de referencia, que nos permita un filtrado previo de los que pueden servirnos para al análisis comparativo.

Posteriormente, tendremos que determinar la necesidad de información complementaria; encuestas, comparación de datos, etc. que se pueda requerir para entrar en la comparación de los elementos clave que nos darán información acerca del impacto de nuestro plan, sobre todo indicadores y patrones comunes de respuesta a las actuaciones implementadas.

Toda esta comparación podrá requerir de la comprobación de la significación estadística de la convergencia o divergencia entre los resultados de cada caso, análisis de regresión, reconstrucción de teorías y procesos, etc.

04

CAPÍTULO 4: UNA PROPUESTA METODOLÓGICA A MODO DE CONCLUSIÓN

Las posibilidades anteriores no son las únicas disponibles para proceder a la evaluación de impacto, pero el objetivo último de esta guía es proporcionar una base suficiente a la vez que asequible, a las personas responsables de la evaluación de un plan estratégico para poder entender este tipo de evaluación y poder tomar decisiones prácticas y fundamentadas al respecto.

No cumpliría su objetivo esta guía básica si no hubiéramos limitado las opciones a aquellas que, dentro de lo razonable, nos puedan dar resultados válidos y fiables, dadas las características del programa que vamos a evaluar. Y tampoco lo haría si no optáramos por una metodología de base sobre la que plantear inicialmente una evaluación de impacto, como primer paso para abordar este tipo de evaluaciones sobre las acciones públicas que desarrolla la Administración, independientemente de que, como no puede ser de otro modo, quedará a criterio de la persona evaluadora la selección final que más se ajuste a sus necesidades.

Por lo tanto, y partiendo de la base de que el objetivo de esta serie de guías del IAAP es suministrar una metodología común y asequible a todo centro directivo de la Junta de Andalucía, que permita fomentar este tipo de evaluaciones como fuente de aprendizaje y validación de las políticas públicas, y que el primer foco se ha puesto en la planificación estratégica, nuestra propuesta para abordar las evaluaciones de impacto de planes estratégicos es la siguiente:

Figura 24: Propuesta metodológica de base para la evaluación de impacto de políticas públicas

Esta es una propuesta general, que recoge todos los elementos que se han ido desgranando a lo largo de este libro, pero no quiere decir que sea necesario realizarlos todos. Esto dependerá del programa que vayamos a evaluar, de lo detallado de su estructura, de lo completo que sea en materia de indicadores, de la claridad del planteamiento, etc. En todo caso, veamos de manera más detallada cada uno de ellos.

Análisis o reconstrucción de la lógica del programa: Como ya se señaló anteriormente, lo primero que hay que encontrar es si el programa tiene una lógica solvente. En este paso se trata de comprobar si las relaciones "causa-efecto" son correctas y tienen una base sólida, teniendo en cuenta evidencias científicas, experiencias previas, etc.

Si fuera necesario reconstruir dicha lógica, tomaremos del documento programático todos los elementos que componen dichas lógicas: objetivos, medidas, etc. y los ordenaremos siguiendo uno de los dos modelos que se han propuesto (EML o lógica del programa).

En este punto puede ser recomendable e incluso necesario, ir más allá de la cadena lógica para comprobar si subyacen hipótesis o ideas preconcebidas poco solventes. En este caso, se recomienda la reconstrucción completa de la Teoría del Cambio, que permite profundizar

más en las relaciones causales que se han establecido entre los elementos del programa[15].

Análisis y disección de objetivos: Previamente a analizar o establecer la Línea Base, será necesario comprobar si se cubren todos los aspectos previstos en los objetivos del programa o, por el contrario, alguno ha quedado sin ser atendido, tanto a nivel de actuaciones como a nivel de indicadores. Para ello procederemos a diseccionar los objetivos como se ha mostrado en el capítulo 2.

Aunque este paso se ha establecido después del análisis de la lógica del programa, realmente es casi simultáneo, ya que es necesario conocer las dimensiones de un objetivo para comprobar lo completo de su estructura y de su lógica.

Selección o análisis de indicadores de la Línea Base: Si el programa cuenta con una Línea Base, en el sentido que se le da en este texto[16], procederemos a analizar la pertinencia y suficiencia de estos indicadores para poder medir efectos relacionados con los objetivos que se persiguen.

Si no contara con una Línea Base será necesario componerla seleccionando los indicadores que mejor se ajustan a los objetivos y sus dimensiones. Puede ser que estos indicadores ya estén definidos en el programa, o que sea necesario recurrir a fuentes externas e, incluso, construirlos expresamente.

Selección del contrafactual: Una vez que contamos con los indicadores y que se tiene clara la actuación realizada, procederemos a construir el contrafactual teniendo en cuenta los indicadores de la Línea Base. En este punto se proponen dos de las metodologías mostradas en el Capítulo 3:

> *Dobles Diferencias*[17]: Usaremos este método en el caso de que el programa se haya dirigido a una población beneficiaria definida y sea posible encontrar población de similares características que no haya participado en el mismo, pero hubiera podido participar en él. Como complemento y siempre que sea posible, se recomienda completar este método con una técnica de matching para minimizar la influencia de factores externos.
>
> *Series Temporales*[18]: Usaremos este método cuando nos enfrentemos a un programa de amplia difusión, que afecta a un contexto amplio más que a una población específica. Para poder usarlo debemos contar con indicadores que puedan modelizarse y proyectarse, por lo que será necesario comprobar, en la literatura especializada o con personas expertas en la materia, la posibilidad real de poder recurrir a este método.

Además de lo anterior, este método es recomendable cuando, siendo aplicable, no se pueda contar con los medios necesarios para realizar otros contrafactuales, o cuando se requiera una evaluación rápida de un programa. En general, el uso de series temporales, o de modelos estadísticos predictivos similares, puede ser la forma más adecuada para

---

15. En la sección "Para Saber Más" se puede encontrar un ejemplo de Teoría del Cambio que puede servir de referencia y referencias bibliográficas a las que dirigirse.

16. Conjunto reducido de indicadores que permiten valorar los cambios en la población o en el contexto previstos en los objetivos del programa.

17. En la sección "Para Saber Más" se puede consultar un ejemplo de este tipo de contrafactual.

18. En la sección "Para Saber Más" se puede consultar un ejemplo de este tipo de contrafactual.

que una Administración se inicie en la sistematización del uso de evaluaciones de impacto para sus programas, especialmente en las evaluaciones de impacto *ex-ante*.

Medida del impacto y conclusiones: Una vez calculada la diferencia de valores en los indicadores de la Línea Base, se procederá a analizar dichos valores y a establecer el impacto o el no impacto del programa.

En este sentido, se deberá concluir cual ha sido el impacto, a la luz de los resultados del análisis del contrafactual y de la lógica del programa, determinando en qué medida la acción pública ha supuesto un cambio real en la población a la que se ha dirigido o en el contexto en el que se ha aplicado.

Será en esta fase cuando se puedan encontrar posibles desviaciones, o resultados anormales o colaterales, que puedan señalar efectos no previstos en el programa. Si se detectan, se recomienda su análisis para poder establecer su relación con el programa y completar el entendimiento de su funcionamiento y efectos, especialmente de cara a futuras intervenciones.

PARA SABER MÁS

En esta sección final se ofrecen ejemplos concretos más detallados sobre lo que se ha descrito en los epígrafes anteriores. La intención es que, quien lo desee, pueda profundizar un poco más en aquellos aspectos de las propuestas metodológicas para la evaluación de impacto, que se hacen en este libro.

En concreto, se incorporan fichas con ejemplos y guías para profundizar en las siguientes cuestiones:

- Diseccionando los objetivos: Dimensiones de los objetivos de diversas políticas públicas.
- Características de los indicadores de contexto.
- Proyectando indicadores: La Línea Base de la Estrategia de Impulso del Sector TIC Andalucía 2020.
- La Teoría del Cambio: Aplicación a un programa educativo.
- Ejemplos de contrafactual en políticas sociales y económicas.

Diseccionando los objetivos: Dimensiones de los objetivos de diversas políticas públicas

Estrategia de Bioeconomía de la Junta de Andalucía. Un ejemplo de objetivos multidimensional

Uno de los trabajos que realizó el área de Evaluacion de Políticas Públicas en IAAP, en 2017 -18, fue un asesoramiento para la elaboración y posibilidad de evaluación de la Estrategia de Bioeconomía de la Junta de Andalucía. A continuación se incluye una resumen de uno de los trabajos que se realizo para el grupo encargado de elaborar la Estrategia.

"En primer lugar queremos felicitar por la mejora evidente en estructuración, coherencia y concreción que permite abordar, ahora sí, la posibilidad de evaluación de esta estrategia centrada en los tres objetivos estratégicos seleccionados:

OBJETIVO ESTRATÉGICO 1	OBJETIVO ESTRATÉGICO 2	OBJETIVO ESTRATÉGICO 3
Incrementar la disponibilidad de biomasa sostenible para su aprovechamiento mediante tratamientos innovadores	Aumentar el número de bioindustrias y biorrefinerías en Andalucía	Incrementar los mercados y el consumo de bioproductos y bioenergía en Andalucía

En nuestro punto de vista, el siguiente paso es realizar la búsqueda y selección de los indicadores que permitan evaluar los objetivos estratégicos. Ello permitirá seleccionar de los mismos, una línea base para realizar el seguimiento, evaluación de resultados y la evaluación de impacto. Para esto proponemos la siguiente secuencia de trabajo:

a) En primer lugar, definir las dimensiones en las que enmarcar los indicadores para cada objetivo.

Dada la complejidad de los objetivos, queremos indicar que habrá que realizar la búsqueda tanto en los sistemas internos de la Consejería como (quizás) en el IECA. En todo caso, los elementos que pueden servir de pistas para la búsqueda de los mismos son los elementos claves (dimensiones) que se están manejando en los propios objetivos y se proponen los siguientes:

Objetivo 1 (4 dimensiones):

- Disponibilidad de biomasa (producción de nueva biomasa + disponibilidad de la ya existente) y grado de aprovechamiento.
- Sostenibilidad de la producción de biomasa.
- Infraestructuras de almacenamiento y transporte (nº, capacidad, etc.).
- Innovación de procesos (proyectos de I+D, grupos de investigación, nº de procesos disponibles, etc.).

Objetivo 2 (2 dimensiones):

- Tamaño del "subsector" de las bioindustrias: volumen y tipo de biomasa procesada, volumen y tipo de bioproductos producidos, n.º de bioindustrias, etc.
- Tamaño del "subsector" de las biorrefinerías: volumen y tipo de bio-

masa procesada, volumen y tipo de bioproductos producidos, nº de biorrefinerías, etc.

Objetivo 3 (3 dimensiones):

- Comercialización de bioproductos (volumen de ventas por tipo de bioproductos, área, etc.).
- Consumo de bioenergía (volumen de ventas por tipo, área, etc.).
- Mercados de productos (transacciones por tipos de mercados: interior/exterior, consumos intermedios o finales, etc.).

En este punto sería necesario que el Servicio de Estadística, junto con el Equipo de Redacción, validen dichas dimensiones (determinar si definen bien los objetivos y si no fuera así proponer cambios) y, a partir de esto, establezcan qué datos/indicadores responden a cada ámbito.Una vez, avanzado este punto, el Servicio de Estadística podrá plantear qué datos/indicadores tiene, cuáles pueden construir en su ámbito de actuación y cuáles considera que deben buscarse en otro lado: IECA, otras Consejerías, Ministerio, etc. separando de cuáles se tienen series históricas (al menos 5 años) y de cuales no y teniendo en cuenta que deben mantenerse en el tiempo.

Por último, una vez determinados estos datos/indicadores, el IAAP formulará la Línea de Base con toda la información anterior, estableciendo las posibilidades de hacer las proyecciones y validando la propuesta tanto con el Servicio de Estadística y como con el Equipo de Redacción de la estrategia.

Finalmente, se harán los cálculos y se planteará el seguimiento de dicha Línea de Base y, si procede, las metas del Plan.

En definitiva, el primer paso es validar la propuesta de dimensiones e ir buscando aquellos datos que ya tengamos disponibles y determinar si es necesario implementar una nueva fuente de información para garantizar que puede incorporarse algún/os indicador/es más.

Equipo de Evaluación de Políticas Públicas del IAAP, 26 de enero de 2018

Características de los indicadores de contexto

¿Qué es un indicador de contexto?

Los indicadores de contexto se definen como aquellos indicadores que proveen de información del entorno social, económico, ambiental, laboral, educativo, tecnológico, etc. aplicados al territorio y/o población. Pretenden identificar la presencia de un fenómeno en un momento y un espacio determinado, siendo una base fiable para establecer comparaciones, bien en el tiempo o bien en el espacio.

Se pueden diferenciar dos tipos de indicadores de contexto:

Indicadores de contexto generales: utilizados para describir la evolución del entorno donde se va a aplicar la intervención pública.

Indicadores de contexto estratégicos (vinculados a los objetivos estratégicos definidos en la intervención) (Herrera M. 2017:3): proporcionan información sobre aspectos relevantes en *relación con la visión* establecida en la Estrategia, Plan o Programa. Relacionan la contribución de la Estrategia/Plan con los *cambios que se irán experimentando* durante el periodo de vigencia, teniendo en cuenta que el contexto se verá también influido por otros factores del entorno ajenos a ella.

Estos indicadores *disponen de un valor base de referencia* (o punto de partida) correspondiente al último dato disponible antes de poner en marcha la intervención, con objeto de comparar con mediciones posteriores y valorar la magnitud de los cambios logrados por la implementación de la intervención pública en aras de realizar evaluaciones de impacto.

Criterios que debe seguir un indicador de contexto

Los objetivos deben ser evaluables, por lo que surge la necesidad de establecer indicadores de contexto que sirvan de punto de partida.

En concreto, los indicadores de contexto deben contar con una cuantificación estandarizada y fiable que simplifique la información, y con ello su comprensión, y permita la realización de comparaciones temporales y/o geográficas. Con ello, y siguiendo los criterios establecidos por la Comisión Europea, la selección de indicadores de contexto debe responder a los criterios *RACER*[19], acrónimo de 5 elementos que se traducen como: *pertinente, aceptado, creíble, fácil de controlar y robusto*.

19. European Commission's Impact Assessment Guidelines. (European Commission, 2005)

Características de los Indicadores de Contexto

Fuentes de información

Las fuentes de las que procedan los indicadores de contexto deben ser fiables y periódicas, para permitir una continuidad en los análisis de la evolución temporal. Por esta razón se dará prioridad a la información procedente de fuentes de información oficiales que permitan un seguimiento continuo que contengan datos regionalizados en función de la realidad que vaya a estudiar.

Algunas de las fuentes de información principales que se proponen son: las oficinas estadísticas regionales, nacionales y europeas. En todas ellas, se establece una diferenciación por temáticas que pasan por cuentas económicas, sectores, condiciones de vida, mercado laboral, empresas, ciencia y tecnología, medio ambientes, etc.

Recoge toda la información de la estadística generada por el Sistema Estadístico y Cartográfico de Andalucía (SECA). https://www.juntadeandalucia.es/institutodeestadisticaycartografia

Se encarga de las funciones de coordinación general de los servicios estadísticos de la Administración General del Estado, la vigilancia, control y supervisión de las competencias de carácter técnico de los servicios estadísticos estataleshttp://www.ine.es

eurostat

Ofrece estadísticas a nivel europeo, con el objeto de realizar comparaciones armonizadas entre países y en el ámbito de regiones NUTS-2 (CCAA) de los países miembro de la Unión Europea. https://ec.europa.eu/eurostat/web/main/home

Ejemplos de utilización de indicadores de contexto en una acción pública de ámbito andaluz

Se toma como ejemplo la Estrategia de Investigación e Innovación para la Especialización Inteligente de Andalucía, RIS3 ANDALUCIA. La elaboración de esta Estrategia es un reto que la Comisión Europea lanzó a todas las regiones con el objetivo final de propiciar un nuevo modelo económico, centrado en las empresas, y basado en una apuesta firme y decidida por la innovación, la ciencia, la tecnología, la internacionalización y la formación.

Los indicadores de contexto que recoge la estrategia son los siguientes:

Visión de Andalucía	Indicadores	Unidad de medida	Fuentes
Andalucía Activa	Producto Interior Bruto	Millones de euros	IECA, INE y Eurostat
		Euros por habitante	IECA, INE y Eurostat
	Ocupados por sexo	Miles de personas	IECA, INE y Eurostat
		Tasa de empleo	IECA, INE y Eurostat
Andalucía Avanzada	Gastos I+D/PIB	Porcentaje	IECA y Eurostat
	Personas con estudios superiores	Número de personas	IECA y Eurostat
Andalucía Abierta	Exportaciones	Miles de euros	Extenda
		Por centaje sobre el PIB	Extenda
	Inversión extranjera directa	Miles de euros	Extenda
Andalucía Saludable	Esperanza de vida al nacer	Años	IECA y Eurostat
	Turistas	Número de personas	IECA
Andalucía Sostenible	Emisiones de gases efecto invernadero	Kilotoneladas equivalentes de Co2	Mº Transición Ecológica y Eurostat
	Consumo energía primaria/PIB	Millones toneladas equivalentes petróleo	Agencia Andaluza Energía y Eurostat
Andalucía Social	Tasa riesgo pobreza o exclusión social	Porcentaje	IECA y Eurostat
	Personas que utilizan TIC para relacionarse con servicios públicos	Porcentaje	IECA y Eurostat

Proyectando indicadores: La Línea Base de la Estrategia de Impulso del Sector TIC Andalucía 2020

Contexto

La crisis económica y financiera acaecida en los últimos años ha puesto de manifiesto las graves carencias estructurales de nuestro modelo actual de desarrollo económico y social. Ante esta situación, tanto la Unión Europea como la Agenda Digital para España se marcan como objetivos mejorar la competitividad del tejido productivo y fomentar su crecimiento, la expansión internacional y la creación de empleo de calidad mediante un mejor aprovechamiento de las TIC y el desarrollo de la Economía Digital, así como impulsar el sistema de I+D+i en TIC. Esta situación cobra especial relevancia en el caso de Andalucía, cuya sociedad se enfrenta a una serie de retos de gran envergadura entre los que se encuentra la transformación de su actual modelo de desarrollo económico. En este contexto, se elabora la Estrategia de Impulso del Sector TIC Andalucía 2020.

Objetivo

La Estrategia de Impulso del Sector TIC Andalucía 2020 tiene por objetivo establecer un planteamiento estratégico que favorezca el desarrollo y consolidación del sector de las Tecnologías de la Información y la Comunicación en nuestra Comunidad Autónoma.

Tipo de estudio e intervención

Esta Estrategia se planteó la consecución de los siguientes objetivos estratégicos:

- *Incrementar la aportación del sector TIC al PIB andaluz.*
- *Aumentar el empleo TIC.*
- *Mejorar la competitividad del sector TIC andaluz.*
- *Estimular la innovación en el sector TIC andaluz.*
- *Favorecer la apertura del sector TIC a nuevos mercados.*
- *Contribuir al desarrollo de la Economía Digital en Andalucía.*
- *Mejorar las condiciones de contexto de desarrollo del sector TIC andaluz.*

Para lograr estos objetivos, establecieron 9 ejes estratégicos, a saber:

- *Gestión del capital humano.*
- *Emprendimiento.*
- *Innovación y especialización tecnológica.*
- *Nuevos mercados y atracción de inversión.*
- *Mejora de la gestión empresarial.*
- *Vertebración del sector TIC.*
- Incorporación de las TIC a otros sectores productivos.
- Financiación empresarial.
- Administración tractora y dinamizadora.

Cómo se hizo

La elaboración de la Línea Base se llevó a cabo en tres fases: recepción y análisis de la información, análisis de los datos y elección del modelo.

- **Recepción y análisis de la información**: Se partió de la disponibilidad de una serie de indicadores que debían cumplir unas características básicas de homogeneización (forma de recogida, escalas utilizadas, periodicidad constante, etc.).
  Para disponer de la información, se llevó a cabo una serie de reuniones con el equipo técnico del Instituto de Estadística y Cartografía de Andalu-

cía (IECA) el cual ha sido el que ha suministrado la información.

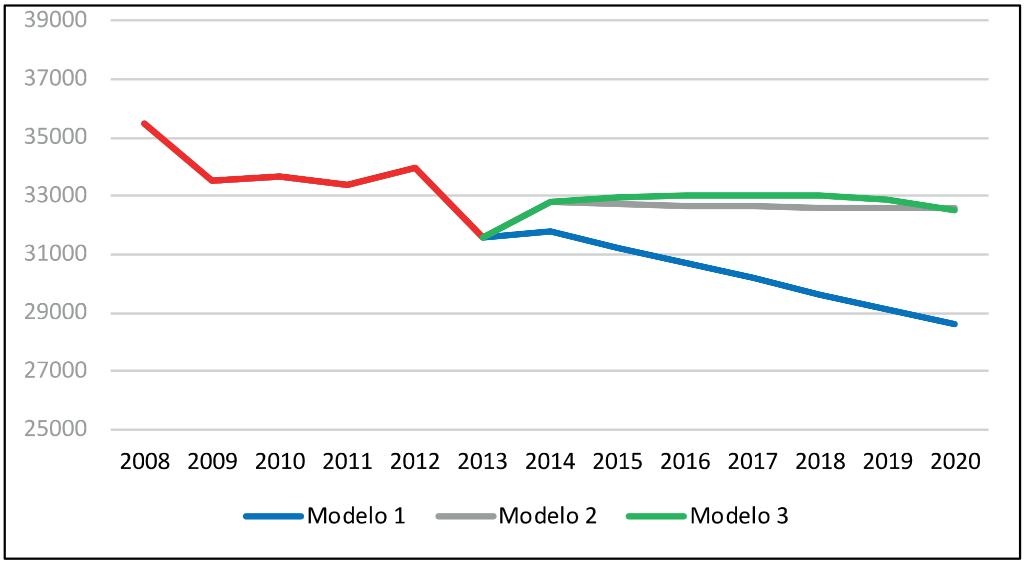
- **Análisis de los datos**: En esta fase, y partiendo de la información recopilada, se realizaron diferentes tipos de análisis en función de cada objetivo planteado.
  - Descripción de los indicadores seleccionados: se realizó un análisis preliminar de la serie que sirviera como base fundamental de la serie temporal, lo que permitió detectar las características más importantes de una serie (tendencia, ciclos, outliers[20]…).
  - Estudio de la tendencia esperable de los indicadores: en función de la tendencia, estacionalidad y dependencia se implementaron los métodos analíticos más exhaustivos y/o que mejor se adaptan al estudio, entre los que podían estar **métodos clásicos** (métodos de descomposición estacional y ajuste de tendencia o métodos de suavizado), **modelos ARIMA** (modelos de medias móviles, autorregresivos, autorregresivos de medias móviles o modelos no estacionarios) y **modelos de regresión** (lineal, logarítmico, inverso, cuadrático, cúbico, potencia, compuesto, curva-S, logística, crecimiento o exponencial).
  - El conjunto de indicadores debe conformar una serie temporal: "Una serie temporal es un conjunto de observaciones que se obtiene midiendo una variable única de manera regular a lo largo de un periodo de tiempo."
- **Elección del modelo**: Para predecir la evolución de cada indicador, se utilizaron tres modelos diferentes, seleccionando de entre ellos aquél cuya probabilidad de evidenciarse fuera más probable. El modelo concreto a seleccionar dependerá en todo caso de cada indicador.

En este sentido, de modo ilustrativo: para el indicador *Peso del Valor Añadido del sector TIC sobre el Producto Interior Bruto en Andalucía*; se realizó un pronóstico de la evolución del Peso del VAB del sector TIC sobre el PIB para Andalucía. Para ello se implementaron tres modelos que han dado como resultado tres escenarios diferentes:

- Modelo 1, calculado con los datos brutos de la serie y mediante mínimos cuadrados; describe un comportamiento parecido al mantenido por los datos hasta el momento de elaboración, una suave caída del Peso del VAB del sector TIC sobre el PIB en Andalucía en 2014 para luego describir un periodo de decrecimiento más pronunciado hasta el 2020.
- Modelo 2, calculado con los datos suavizados de la serie y mediante regresión cúbica; describe un modelo que mantiene la tendencia decreciente hasta el año 2017, momento en el que comenzó una tendencia creciente del Peso del VAB del sector TIC sobre el PIB en Andalucía y que aumentaría hasta 2020.
- Modelo 3, calculado con los datos brutos de la serie y mediante regresión cuadrática, describe el modelo más pesimista de los estimados, es una continuación de la tendencia decreciente de los datos que se disponían hasta el momento.

20. Outlier: un elemento de los datos que es significativamente diferente a los otros datos de la colección, o un elemento que parece implicar un patrón que es inconsistente con el grueso de la evidencia de datos.

Modelos para el Número de ocupados del sector TIC en Andalucía-2020



A partir de la información reflejada por los modelos, se obtendrían conclusiones que deberían permitir al equipo redactor del documento de planificación determinar cuál de los tres modelos es el mejor para ajustarse al escenario futuro que se evalúa.

La Teoría del Cambio: Aplicación a un programa educativo.

¿Qué es la teoría del cambio?

Se pueden encontrar distintas definiciones sobre la Teoría del Cambio, si bien todas tienen en común que se trata de un enfoque de pensamiento y acción, un proceso de reflexión que trata de explorar el cambio y cómo sucede, haciendo explícitas y articulando las suposiciones subyacentes a cómo se cree que el cambio sucede. En definitiva, se trata de un proceso que tiene por objetivo explicitar las hipótesis, ideas y asunciones sobre cómo se supone que debe funcionar el programa o política que se está analizando, mostrando los fundamentos lógicos sobre los que se estructura y gestionan los mismos.

¿Para qué se usa la teoría del cambio?

La Teoría del Cambio, además de ser una herramienta en sí misma, se puede utilizar principalmente para tres usos:

a) **Diseño del programa**. Elaborar una Teoría del Cambio antes de iniciar cualquier tipo de intervención puede llevar a extraer importantes conclusiones antes de su puesta en práctica. Así, se puede detectar si se está haciendo un uso racional de los recursos (humanos, económicos, logísticos, etc), si la lógica de la intervención está correctamente definida, si se dan los elementos necesarios para la correcta alineación de las estrategias de relación que se determinan entre los diferentes actores para apoyar una acción y, por último, si se ha contemplado el marco para el seguimiento y evaluación de la política o programa.

b) **Rediseño o ajuste de las intervenciones.** durante la fase de ejecución de la política o programa. La Teoría del Cambio se puede utilizar para revisar la totalidad de la política o programa, así como partes concretas de las mismas durante el período de ejecución. La utilización de esta herramienta pude servir para mejorar, descartar o incorporar nuevas actividades o para sacar a la luz los supuestos subyacentes para desecharlos, corregirlos o elaborar otros nuevos.

c) **Evaluación**. La definición de una Teoría del Cambio puede servir como base para una evaluación, ya que proporcionará las preguntas de evaluación, descubrirá los supuestos críticos a ser explorados, identificará los resultados a medio y largo plazo y sus vínculos causales, a la vez que definirá los indicadores clave a ser medidos y/o analizados. Determinará de esta forma, la valoración y pertinencia de la intervención y las recomendaciones para posteriores fases o nuevas intervenciones.

Siguiendo a Ligero Lasa (2011), la evaluación de la Teoría del Cambio tiene como finalidad comprender las cadenas o secuencias causales que provocan los resultados que se esperan conseguir con una política o programa. Su intención, por tanto, es saber por qué han ocurrido las cosas y en su caso, mejorar lo que se haya detectado como mejorable en la intervención. De esta forma, se puede afirmar que su finalidad principal es la "mejor comprensión" de la política o programa.

¿Cómo se desarrolla la teoría del cambio? Pasos básicos[21]

Los pasos básicos para el desarrollo de la Teoría del Cambio y las preguntas

21. Extracto del ejemplo práctico desarrollado en el curso "Teoría del Cambio" impartido por Daniel Guijarro para el IAAP en mayo de 2018. Grupo de trabajo formado por Adolfina Martínez Guirado, Sofia González Gómez, Antonio Feria Moreno y Andrés Márquez Vázquez.

orientadoras para configurar cada uno de estos pasos son:

1. **Formular el cambio deseado:**
   ¿Qué cambio se desea que ocurra?
   ¿Cuál es el propósito del cambio al que se quiere contribuir?
   ¿Cuáles son las cuestiones que se quiere ayudar a cambiar?

2. **Analizar el ecosistema, tanto político, como social, cultural, histórico, económico, etc.:**
   ¿Cuáles son las condiciones políticas, sociales, históricas y económicas que afectan o son afectadas por el proceso de cambio?
   ¿Qué oportunidades se detectan en el contexto?
   ¿Cuáles son los actores involucrados en el proceso de cambio y cómo se relacionan entre sí?
   ¿Con qué intervalo de tiempo se está trabajando?

3. **Definir las condiciones para el cambio.**
   ¿Qué necesita estar en su lugar para que el cambio deseado llegue a ocurrir? (marco jurídico, financiación, instituciones)
   ¿Cuáles son los obstáculos para que estas condiciones no estén en su lugar?
   ¿Cuáles son las oportunidades para que el contexto ofrece para que estas condiciones ocurran?

4. **Identificar constantemente los supuestos de partida y los procesos:**
   ¿En qué supuestos iniciales se basa el cambio deseado?
   ¿Por qué se piensa que las intervenciones contribuirán efectivamente a esas condiciones y al cambio deseado?
   ¿Con quién se han compartido y discutido estos supuestos?

5. **Definir la contribución al cambio (contribución a la condición en concreto):**
   ¿Cuál es la contribución más efectiva que se puede aportar?
   ¿Cómo se relaciona la contribución propuesta con otras contribuciones procedentes de diferentes actores?
   ¿Qué se puede hacer para aumentar la probabilidad de que se produzcan las condiciones para el cambio?

6. **Analizar los actores del cambio:**
   ¿Quiénes son las partes interesadas clave que afectan o se ven afectadas por el proceso de cambio?
   ¿Cuáles son los intereses y las posiciones de estas partes interesadas en relación con el proceso?
   ¿Cuáles son los supuestos subyacentes de las partes interesadas, sus relaciones y alianzas estratégicas?

7. **Definir las Estrategias de ejecución e implementación:**
   ¿Qué estrategias se necesita implementar para actuar con eficacia?
   ¿Qué estrategias se necesita implementar para mejorar las relaciones entre las partes interesadas?

8. **Desarrollar y definir los indicadores que cuantificarán el cambio, los logros y los progresos:**
   ¿Quién decide lo que se necesita observar para saber si hemos logrado el cambio deseado?
   ¿Cómo se va a recoger la información que indica el cambio?
   ¿Qué resultados no esperados y/o negativos han tenido lugar debido a las acciones desarrolladas?

Un ejemplo práctico

A continuación, se desarrolla un ejemplo práctico de la Teoría del cambio en programas que se están ejecutando actualmente.

Programa para el desarrollo de la competencia en comunicación lingüística

http://portal.ced.junta-andalucia.es/educacion/webportal/documents/343370/0/Programa+para+el+desarrollo+de+la+competencia+en+comunicaci%C3%B3n+ling%C3%BC%C3%ADstica.+Marco+de+referencia.pdf

Cambio deseado: Debe ser específico y al menos debe responder a quién es el sujeto del cambio y qué se pretende conseguir.

El alumnado en Andalucía mejora su formación personal, la adquisición de conocimientos, la autonomía para aprender y el desarrollo integral como persona.

Condiciones para el cambio: Situaciones que se tienen que dar para que el cambio deseado suceda en realidad, teniendo en cuenta el análisis de contexto realizado.

Condición 1: El profesorado dispone de los recursos y actividades susceptibles de ser incluidos en las programaciones didácticas

Condición 2: Las instituciones públicas y fundaciones de la región como encargadas de la difusión del legado cultural y literario facilitan la realización de actividades propuestas en el aula.

Supuestos para el cambio: Un supuesto es aquello que se toma como cierto sin cuestionar su veracidad. Su valor parte de que hace explícitos los puntos de vista y suposiciones de los involucrados en el proceso. Se trata de descubrir las suposiciones que se tienen sobre cómo se cree que el cambio puede llegar a suceder, haciendo la lógica de intervención coherente. Los supuestos están directamente relacionados con las condiciones para el cambio.

Supuesto para la Condición 1: La disposición de medios y recursos, así como la mejora metodológica por parte del profesorado facilita su labor a la hora de realizar las programaciones didácticas y permiten que el alumnado desarrolle su competencia comunicativa.

Supuesto para la Condición 2: La colaboración de las instituciones culturales y educativas pone a disposición del profesorado un mayor número de recursos para su uso en el aula.

Contribuciones a la Condición (CC): Se entiende como aquellas estrategias que el programa quiere poner en marcha para aumentar las posibilidades de que las condiciones formuladas se hagan realidad. Y con ello estarán contribuyendo al cambio que quieren ver.

Contribución a la Condición 1: Aplicar una metodología participativa basada en trabajo cooperativo y en las unidades didácticas por proyectos, utilizando la herramienta "Cuaderno de propuestas Estratégicas".

Contribución a la Condición 2: Puesta en marcha de un calendario de visitas trimestral de los centros docentes a los centros/instituciones culturales y educativas de las distintas poblaciones.

Indicadores: Miden el cambio que se desea alcanzar.

Indicador: Mejora de competencia lingüística del alumnado de los centros educativos (PISA)

Teoría del Cambio del programa para el desarrollo de la comunicación lingüística

EJEMPLOS DE CONTRAFACTUAL EN POLÍTICAS SOCIALES Y ECONÓMICAS

Ejemplo 1: Contrafactual basado en grupos de tratamiento y control

Impacto de la iniciativa de cooperación social y comunitaria para el impulso del empleo de mayores de 30 años[22].

CONTEXTO

La crisis económica y financiera cuya mayor incidencia se dio en Andalucía a partir de 2011, acarreó efectos devastadores en el mercado laboral de la Región. El conjunto de personas con 30 o más años se identificó como uno de los más afectados por la pérdida de empleo, unido a las mayores dificultades para su reincorporación al mercado laboral. Es por ello por lo que, desde los organismos competentes en materia de políticas activas de empleo, se decidió la puesta en marcha de una iniciativa generadora de empleo dirigida a la población de treinta o más años de edad desempleada. Esta iniciativa, denominada Iniciativa de Cooperación Social y Comunitaria 30+, establece, además, dentro del colectivo de personas destinatarias, un orden de prioridad en función de su tiempo de permanencia en situación de desempleo y de la situación de especial vulnerabilidad en la que se encuentren.

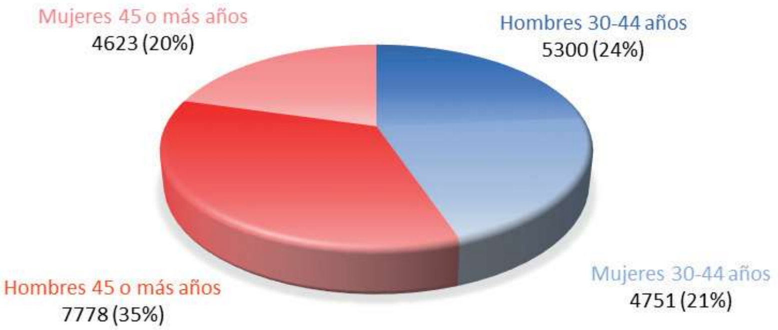
OBJETIVO

La Iniciativa Cooperación Social y Comunitaria para el impulso del empleo 30+, (cofinanciada por el P.O. FSE de Andalucía 2007-2013) tiene por **objeto promover la mejora de la empleabilidad de las personas de treinta o más años de edad desempleadas, incrementando su experiencia profesional**, a través de la contratación por parte de los ayuntamientos andaluces para el desarrollo de proyectos de utilidad colectiva que redunde en beneficio de la comunidad, impulsando la adquisición o recuperación de competencias profesionales de las personas contratadas para su integración en el mercado laboral.

TIPO DE ESTUDIO E INTERVENCIÓN

Con el análisis de las contrataciones realizadas por parte de los ayuntamientos andaluces receptores de las ayudas, se constató que el grado de penetración del programa en la población desempleada andaluza de más de 30 años fue del 2.9%. Esto es, de cada 100 desempleados registrados en el Servicio Andaluz de Empleo (SAE) de más de 30 años en Andalucía, se beneficiaron del programa 2.9. La distribución por sexo y edad fue la que sigue:

22. Evaluación incluida en el en el Informe de Evaluación Operativa por Reprogramación P.O. FSE de Andalucía 2014-2020, elaborado por la Dirección General de Fondos Europeos con la Asistencia Técnica de SAET Consultores y la Colaboración del Dr. José María Millán Tapia [Universidad de Huelva]. Junio de 2018.

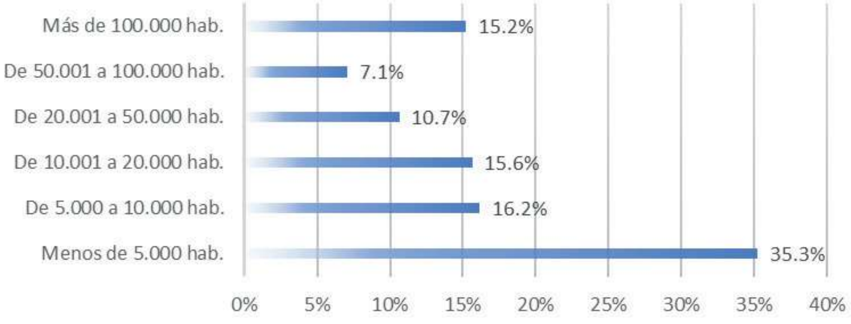


Distribución por sexo y edad de los beneficiarios

A nivel municipal, se beneficiaron de este programa desempleados del 97.6% de los municipios andaluces (759 de los 778), lo que era coherente con el interés prioritario de que esta iniciativa se desarrollara en todos los municipios del territorio, siendo la distribución por provincias la que sigue:

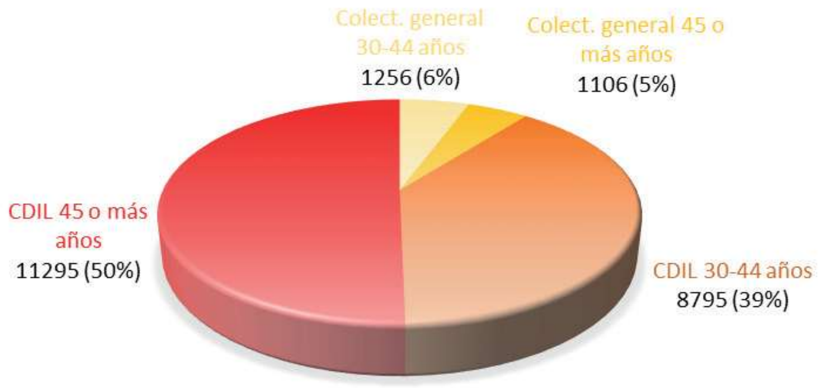
Número de beneficiarios en las provincias andaluzas

En términos más *cualitativos*, se observó cómo más de la mitad de los beneficiarios de este programa lo fueron en municipios de menos de 10.000 habitantes, donde las capacidades para el desarrollo económico son más reducidas.



Distribución de beneficiarios según tamaño del municipio

En esa misma línea, el 55% de las beneficiarias lo conformaron personas con 45 o más años, cuyas dificultades a la hora de encontrar un empleo son mayores. Además, se comprobó como del total de personas beneficiarias, el 89% pertenecía a colectivos de difícil inserción laboral (CDIL) conformados mayoritariamente por parados de larga duración, personas sin estudios secundarios, personas sin experiencia profesional y beneficiarios del Ingreso Mínimo de Solidaridad, aunque también se incluía a otros colectivos vulnerables como personas con diversidad funcional, inmigrantes etc. En este sentido, más de la mitad de las personas beneficiarias del programa fueron personas de 45 o más años pertenecientes a al colectivo de difícil inserción (CDIL), y 1 de cada 6 beneficiarios fueron personas con 45 o más años caracterizadas simultáneamente por ser parados de larga duración y por carecer de estudios secundarios.



EFECTOS DEL PROGRAMA

Del total de beneficiarios del programa, un total de 5.379 obtuvieron un contrato en los 6 meses posteriores a la finalización del programa, lo que depara una **tasa de contratación posterior del 24%.** Junto a estos beneficiarios con contrato posterior, otra medida de empleabilidad posterior del programa fue la inscripción en el Registro de Trabajadores Autónomos de 172 beneficiarios a los 6 meses de la fecha de su finalización, lo que genera una tasa global de empleabilidad posterior para el **conjunto de beneficiarios del programa del 24.7%.** Dicho de otro modo, 1 de cada 4 participantes en este programa estaba empleado a los 6 meses de finalización del mismo.

RESULTADOS DE LA EVALUACIÓN

La pregunta formulada fue: *¿qué les hubiese pasado a los desempleados participantes en la Iniciativa si no hubieran participado en la misma?*

Para poder valorar la diferencia entre el resultado observado y el "contrafactual" se compara la situación de dos grupos de individuos, uno de los cuáles ha participado en el programa (conocido como *grupo de tratamiento* o simplemente *tratados*) y el otro no (conocido como *grupo de control* o simplemente *controles*).

21.255
Grupo Tratamiento

760.788
Grupo Control

Para este análisis en concreto se utilizó la técnica de *matching* o emparejamiento y, en concreto, el *Propensity Score Matching* (PSM) o **método de emparejamiento basado en la probabilidad de asignación**.

En esta evaluación el grupo de tratamiento lo formaron los participantes en el Programa Emple@ 30+ pertenecientes a la convocatoria que es objeto de análisis. Por su parte, el grupo de control lo formaron todas las personas desempleadas de más de 30 años dadas de alta en el SAE en mayo de 2015, fecha fin del programa (excluido de este grupo los participantes en el Programa Emple@ 30+).

Modelo general	*Propensity Score Matching*
Método específico de estimación	Nearest Neighbour (se obtienen resultados similares mediante otros métodos como el método Kernel o del núcleo)
Variable de efecto	Probabilidad de obtener un contrato en los 6 meses siguientes a la finalización de la iniciativa (para los tratados; mayoritariamente mayo de 2015) y en los 6 meses siguientes a mayo de 2015 (para los controles)
Variable de tratamiento	Participación o no en el Programa Emple@ 30+
Variables independientes	Género, edad, nivel educativo, duración en el desempleo, ser beneficiario o no del IMS, experiencia previa en el mercado laboral, ser persona con diversidad funcional, ser extranjero o no

El resultado del análisis puso de manifiesto que el Programa Emple@ 30+ generó un aumento de la empleabilidad para el colectivo de tratados con respecto a los controles en torno a un 3%, lo que expresado como porcentaje (%) de la tasa de empleabilidad de los tratados, se traduce en un aumento relativo del 12,6%. **Es decir, la probabilidad de ser contratado tras la participación en la iniciativa aumenta un 12,6% respecto a los no participantes**.

Ejemplo 2: Contrafactual basado en series temporales

El impacto de la política monetaria no convencional sobre las finanzas públicas de la UEM[23]

CONTEXTO:

El Banco Central Europeo (BCE) adoptó en junio de 2014 una serie de **medidas de política monetaria que permitieron normalizar las condiciones financieras en el área del euro** y restablecer el buen funcionamiento de los mecanismos de transmisión monetarios ante un contexto de crisis financiera y económica.

Las actuaciones llevadas a cabo fueron: fijar el tipo de interés de la facilidad de depósito hasta el –0,4 %, además de acentuar el impulso monetario a través de la expansión y la recomposición de su balance, con «Operaciones de Financiación a Largo Plazo con Objetivo Específico», para proveer financiación a largo plazo al sistema bancario y reforzar la transmisión de la política monetaria a través del canal crediticio, y por último, en diciembre de 2014, introdujo un programa de compras de activos a gran escala.

OBJETIVO:

Las actuaciones de política monetaria impulsadas por el BCE han sido eficaces para moderar las condiciones de financiación de la economía, contribuyendo al fuerte descenso observado en los tipos de interés a los que se financia la deuda soberana, directamente relacionado con los pagos por intereses asociados a dicha deuda.

Con estudio se ofrece una cuantificación de los efectos sobre la situación de las finanzas públicas de la Unión Económica y Monetaria (en adelante, UEM) y de algunos de sus principales países de referencia, como España. En concreto, se proporcionan estimaciones del efecto directo de las medidas monetarias sobre el pago de intereses de la deuda soberana[24].

TIPO DE ESTUDIO E INTERVENCIÓN:

Para el análisis de las medidas de política monetaria no convencional se propone un estudio de proyecciones, basada en los ejercicios de proyección del Eurosistema, dónde las previsiones se obtienen a partir de tipos *forward*. Estas previsiones se comparan con la senda real observada.

Junto a ello, se aplica el denominado "estudio de eventos"[25] en el que se determina la hipótesis de que en uno o dos días las variaciones de los activos financieros responderán sólo a esos cambios, utilizando esa evolución para aislar efectos.

Por último, se relaciona la trayectoria observada de la carga de intereses con la producida si los tipos de interés de emisión de la deuda a corto (un año)

23. BURRIEL, P., MARTÍ, F. y PÉREZ, J.J. (2017) Impacto de la política monetaria no convencional sobre las finanzas públicas. Madrid. Boletín Económico 3/2017. Banco de España

24. Asimismo, el estudio presenta una estimación aproximada de los efectos totales (directos e indirectos) de la política monetaria no convencional sobre las finanzas públicas para España y el conjunto de la UEM.

25. Se consideran unos 60 eventos, 20 de los cuales se refieren a conferencias de prensa del BCE, 13 a publicaciones de las deliberaciones del Consejo de Gobierno y el resto a intervenciones públicas de miembros del Consejo de Gobierno del BCE.

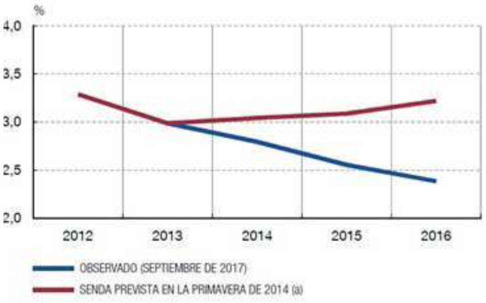
y largo plazo (cinco años) se hubieran situado en los niveles previstos antes de las medidas de política monetaria no convencional, teniendo en cuenta los ejercicios de proyección del eurosistema, a través de la curva de tipos *forward*, y con el supuesto no variación en la situación presupuestaria.

RESULTADOS DE LA EVALUACIÓN:

La **proyección** referida mostraba una estabilización del **tipo implícito de la deuda** en los niveles vigentes en aquel momento, con una previsión coherente con los tipos de interés de emisión medidos por los tipos *forward* implícitos en la curva de tipo estimada en mayo de 2014 para los plazos señalados, que señalaba que el tipo se situara en el 3,2% en el conjunto del área del euro y en el 3,6% en España, en 2016.

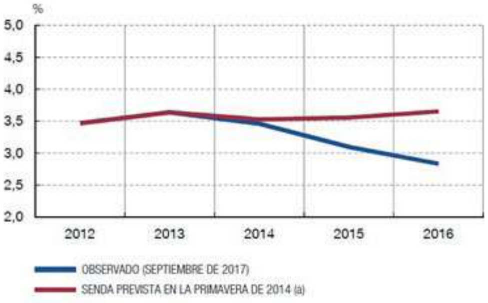
Estas proyecciones realizadas se han comparado con los resultados reales concluyendo que ha habido una reducción de los tipos, siendo el valor finalmente observado en la UEM del 2,4% y en España del 2,8 %, por encima del 3,2% y 3,6%, comentando anteriormente.

Tipo implícito de la deuda pública. UEM

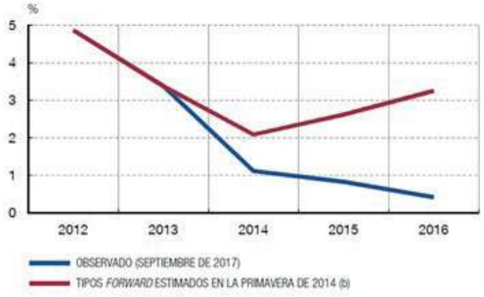


Tipo de emisión de la deuda pública a cinco años. UEM

Tipo implícito de la deuda pública. España



Tipo de emisión de la deuda pública a cinco años. España



Además, el estudio propone para aislar efectos no atribuibles a la política monetaria, quedarse sólo con los ocurridos como consecuencia de las actuaciones sobre el precio de los activos financieros, y al objeto de estudiar la senda de la rentabilidad de la deuda pública en la UEM en el entorno cuando se realizaron los anuncios de medidas de política monetaria por parte del BCE, se determina

la hipótesis de que en uno o dos días las variaciones de los activos financieros responderán sólo a esos cambio. Esta hipótesis es conocida como el "estudio de eventos"[26]. Con ello, se podría atribuir a la política monetaria cerca del 75 % de la caída observada en los tipos de interés a cinco años desde junio de 2014 hasta finales de 2016, en la UEM.

Junto a ello, y utilizando las sendas descritas, se cuantifica el efecto de esta política sobre los tipos de emisión de la deuda a los distintos plazos relevantes para determinar el gasto en intereses, con una hipótesis de comparación de la carga de intereses de emisión de deuda, con las proyecciones de tipos *forward* antes de la intervención y con una situación presupuestaria (saldo público primario, ajuste déficit-deuda y duración media de la deuda) estable.

Con este análisis, se concluye que, entre 2014 y 2016, fruto de las medidas de política monetaria no convencional, y siendo éste el contrafactual, España se habría ahorrado en torno a un punto porcentual del PIB, en concepto de pagos por tipos de la deuda soberana, siendo el ahorro obtenido por la UEM de medio punto porcentual del PIB.

	Variación del tipo de interés de la deuda pública								Ejercicio contrafactual (a)
	Observada		Debido a la política monetaria no convencional (b)				Esperada en mayo de 2014		Ahorro en gasto de intereses (% del PIB)
			Porcentaje		Variación				
	Un año	Cinco años	Un año	Cinco años	Un año	Cinco años	Un año	Cinco años	
UEM	-34	-108	75	74	-25	-80	90	96	0,5
España	-88	-139	23	84	-20	-117	157	127	0,9

Fuentes: El impacto de la política monetaria no convencional sobre las finanzas públicas de la UEM, según datos de Eurostat y Banco de España

(a) Diferencias con respecto al ejercicio contrafactual consistente en asumir que, en ausencia de la política monetaria no convencional, los tipos de interés habrían evolucionado en línea con lo previsto en mayo de 2014 (tipos *forward* de mercado).

(b) Estudio de eventos para el período junio de 2014-diciembre de 2016, considerando la variación con una ventana de dos días en torno a 36 fechas en las que el Consejo de Gobierno del BCE anunció medidas, se publicaron reseñas de sus reuniones o se produjeron intervenciones de destacados miembros del Consejo que permitían anticipar la posibilidad de la adopción próxima de actuaciones.

26. Esta estimación es una actualización con datos hasta principios de 2017 del estudio de eventos elaborado por Banco de España (2016). En particular, se consideran unos 60 eventos, 20 de los cuales se refieren a conferencias de prensa del BCE, 13 a publicaciones de las deliberaciones del Consejo de Gobierno y el resto a intervenciones públicas de miembros del Consejo de Gobierno del BCE.

Ejemplo 3: Otra forma de medir impactos: efecto antes-después de la intervención

Evaluación ex post de la inversión en Infraestructura de Transporte por Carretera y Ferrocarril apoyadas por el FEDER y el Fondo de Cohesión

CONTEXTO:

El elevado volumen de *inversión pública en el sistema de transporte* en el período analizado 2000-2013 ha llevado a establecer un proceso de convergencia de España con sus economías de referencia, tanto en términos de dotación de stock, como de calidad de las infraestructuras, movilidad y siniestralidad, aunque, a pesar de la evolución favorable, continúan los principales problemas.

Estas inversiones en infraestructuras han sido claves en el dinamismo de la actividad productiva, y por ende, de las mejora en la variables macroeconómicas que afectan a la sociedad.

OBJETIVO:

Con estudio se persigue medir los *efectos en términos de producción y empleo* de las infraestructuras de transportes por carreteras y ferrocarril, tras la implementación de políticas de cohesión en el ámbito nacional.

TIPO DE ESTUDIO E INTERVENCIÓN:

Se aplica un modelo de demanda que considera la inversión pública, de los Fondos FEDER y de Cohesión en los períodos 2000-2006 y 2007-2013 en proyectos de infraestructuras de transporte por carretera y ferrocarril, como *elemento exógeno impulsor* de la actividad económica y de la generación de empleo y renta.

La inversión en infraestructuras, además del efecto directo que tiene sobre la producción y el empleo, cuenta con un elevado efecto multiplicador en otros sectores, que lleva a generar inversiones adicionales en el resto de la economía, y arrastra a la demanda de productos y empleo en otras ramas de actividad para satisfacer la producción y la demanda. Por ello, se decide utilizar un **análisis del marco input-output**, para recoger el total del impacto en producción y empleo, que incluya los efectos directo, indirecto e inducido.

RESULTADOS DE LA EVALUACIÓN:

Para este estudio se requirieron tres fases de trabajo: 1) recopilación y tratamiento de la información, 2) estimación y cálculos previos y 3) modelos lineales para el cálculo de impacto:

Fase 1. Recopilación y tratamiento de la información: se utiliza matriz input-output multiregional[27] y al encontrarnos en una economía abierta, se completa, con las obtenidas para la Unión Europea y el resto del mundo.

Junto a ello, se trabaja en la variable exógena, inversión, identificando tipología (ferrocarril y carreteras) y reparto por Comunidades Autónomas.

Además, se utiliza una matriz de reparto de inversiones obtenida de la clasificación del Ministerio de Fomento y Eurostat.

27. Multiregional Input–Output Model for the Evaluation of Spanish Water Flows. Rosa Duarte e Ignacio Cazcarro. Universidad de Zaragoza. 2013.

Fase 2. Estimación y cálculos previos: dado que la Tabla Input Output (TIO) multirregional data de 2005, se utilizó el método Euro, usado por Eurostat, para obtener la proyección de TIO multirregional en 2006 y 2015. Junto a ello, se utiliza información actualizada de la Contabilidad Regional del INE para completar las estimaciones.

Fase 3. Modelos Lineales para el cálculo de impacto: para el cálculo del impacto desglosado en efectos directos, indirectos e inducidos en términos de producción y empleo se ha utilizado Matriz de Contabilidad Social (MCS) a partir de las TIO.

Estas MCS considera 40 ramas de actividad, cuentas endógenas (sectores productivos y consumidores) y cuentas exógenas (sector público, sector exterior y diferencia entre ahorro e inversión).

En la elección de modelo de multiplicadores, se seleccionan modelos multisectoriales de tipo lineal[28] en el que las variables endógenas se expresan como función lineal de las variables exógenas.

Tras los trabajos previos comentados, los resultados del estudio que se plasman en este ejemplo se van a centrar en los efectos en la producción y el empleo, observando los impactos absolutos y relativos[29].

Los resultados obtenidos muestran que la inversión pública derivada de la ejecución de las obras de carreteras y ferrocarriles en el período evaluado ha provocado un impacto positivo en la ***producción***, llevando a duplicar su valor (ha pasado de 36.180 millones de euros a 66.758 millones de euros.

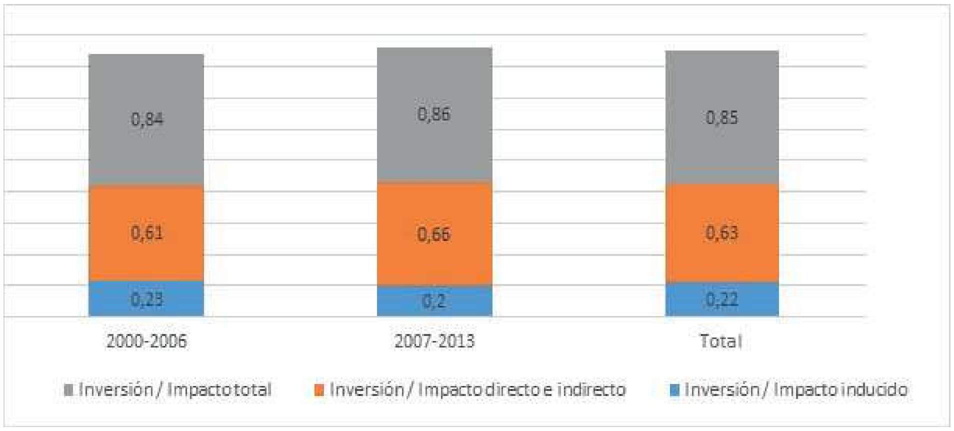
	Inversión inicial	Impacto directo e indirecto	Impacto inducido	Inversión+ Impacto Total
2000-2006	24.939.197	15.204.519	5.691.473	45.835.189
2007-2013	11.241.209	7.409.472	2.271.828	20.922.509
TOTAL	**36.180.406**	**22.613.991**	**7.963.301**	**66.757.698**

Impacto en términos de producción. Valores absolutos en miles de euros

28. Pérez y Pérez, et al, 2011.

29. Como se ha comentado la evaluación de impacto a través de MIO permite obtener análisis diferenciados por sectores y CCAA, se remite al estudio referenciado para ampliar esta información.

Con ello, y en términos relativos de cada euro invertido se han generado 0,85 céntimos de euro adicionales, en promedio, de forma directa e indirecta (0,63) e inducida (0,25) y manteniéndose prácticamente constante en los dos períodos de tiempo analizados.



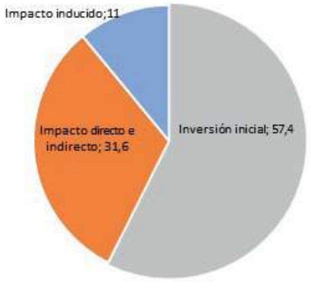
Impacto en términos de producción. Valores relativos

En términos de **empleo**, se han creado más de un millón de empleos en el periodo evaluado, siendo casi medio millón de ellos (428.072) derivados de los efectos directos, indirectos e inducido de la aplicación de la política.

	Empleo derivado de la inversión inicial	Impacto directo e indirecto	Impacto inducido	Impacto Total
2000-2006	399.120	207.474	78.066	684.660
2007-2015	178.770	110.127	32.405	321.301
TOTAL	**577.889**	**317.601**	**110.471**	**1.005.962**

Impacto en términos de empleo. Valores absolutos

El 57,4% del empleo generado se debe a la inversión inicial, correspondiendo el 31,6% al impacto directo e indirecto y el 11% restante al inducido.



Distribución del efecto de la inversión en el empleo por tipo de impacto

BIBLIOGRAFÍA

BLASCO, J. (2009). Guía práctica 3 - Evaluación del diseño. Barcelona: Ivàlua – Institut Català d'Avaluació de Polítiques Públiques.

BLASCO, J., y CASADO, D. (2009). Guía práctica 5 - Evaluación de impacto. Barcelona: Ivàlua – Institut Català d'Avaluació de Polítiques Públiques.

BURRIEL, P., MARTÍ, F. y PÉREZ, J.J. (2017). Impacto de la política monetaria no convencional sobre las finanzas públicas. Madrid. Boletín Económico 3/2017. Banco de España. Consultado en https://www.bde.es/f/webbde/SES/Secciones/Publicaciones/InformesBoletinesRevistas/ArticulosAnaliticos/2017/T3/fich/beaa1703-art26.pdf en diciembre de 2018.

COMISIÓN EUROPEA – EuropeAid. (2001). Manual de gestión del ciclo del proyecto. Consultado en http://www.bantaba.ehu.es/formarse/ficheros/view/Manual_de_Gesti%C3%B3n_del_Ciclo_del_Proyecto.pdf?revision_id=69278&package_id=69185 en septiembre de 2018.

FERIA MORENO, A. (2017). Manual de elaboración de planes estratégicos de políticas públicas en la Junta de Andalucía, Sevilla: Instituto Andaluz de Administración Pública.

GERTLER, Paul J., MARTÍNEZ, Sebastián; PREMAND, Patrick, RAWLINGS, Laura B. y VERMEERSCH Christel M. J. (2017). La evaluación de impacto en la práctica, Segunda edición. Washington, DC: Banco Interamericano de Desarrollo y Banco Mundial.

GOODRICK, D. (2014). Estudios de caso comparativos, Síntesis metodológicas: evaluación de impacto n.º 9, Centro de Investigaciones de UNICEF, Florencia. Consultado en https://www.unicef-irc.org/publications/pdf/MB9ES.pdfen noviembre de 2018.

GUIJARRO, D. (2018). Material del curso de Teoría del Cambio impartido por el IAAP en mayo de 2018.

HERRERA MENCHÉN, M. (2017). Diseño de la evaluación. Sevilla: Instituto Andaluz de Administración Pública.

LIGERO LASA, J. A. (2011). Dos métodos de evaluación: criterios y teoría del programa. Serie CECOD Número 15 / 2011. Madrid: CEU Ediciones.

OECD Directorate for Science, Technology and Innovation (2014). Assessing the impact of state interventions in research – Techniques, issues and solutions y consultado en https://www.oecd.org/sti/inno/What-is-impact-assessment-OECDImpact.pdfen junio de 2018.

PÉREZ Y PÉREZ, L., SANAÚ VILLARROYA, J.J., y SANZ VILLARROYA, I., (2011). Impacto macroeconómico de las inversiones en la red de transporte de la electricidad en España. Estudios de Economía Aplicada (Vol. 29, N.º 2) .

RETOLAZA I. (2011) Teoría de Cambio. Un enfoque de pensamiento y acción para navegar en la complejidad de los procesos de cambio social. Guatemala: PNUD/VIHOS

ROGERS, P. (2014). Sinopsis de la evaluación de Impacto, Síntesis metodológica n.º 1, Centro de Investigaciones de UNICEF, Florencia. Consultado en https://www.unicef-irc.org/publications/pdf/brief_1_overview_spa.pdfen junio de 2018.

Evaluación Políticas Públicas

www.juntadeandalucia.es/
institutodeadministracionpublica

redprofesional.juntadeandalucia.es
Grupo "EVALUACIÓN DE POLÍTICAS PÚBLICAS"

www.scoop.it/t/evaluacion-de-politicas-publicas

evaluacion.politicas.iaap@juntadeandalucia.es

IAAP - Instituto Andaluz de
Administración Pública

@EPP_JdA

Instituto Andaluz de
Administración Pública

JUNTA DE ANDALUCIA
Instituto Andaluz de Administración Pública
CONSEJERÍA DE LA PRESIDENCIA,
ADMINISTRACIÓN PÚBLICA E INTERIOR